

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة البصرة- كلية الإدارة والاقتصاد

قسم الاحصاء

نمذجة تقلبات الحمل الكهربائي والتنبؤ به باستخدام انحدار المتجه الداعم مع

الخوارزميات الفوضوية والمنطقة الجنوبية

رسالة تقدمت بها

هدى عبد السادة هاشم الربيعاوي

الى مجلس كلية الادارة والاقتصاد / جامعة البصرة

وهي جزء من متطلبات نيل درجة الماجستير علوم في الاحصاء

اشراف

أ.م. ساهره حسين زين الثعلبي

2021 م

1443هـ

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

يَكَادُ الْبَرْقُ يَخْطَفُ أَبْصَارَهُمْ كُلَّمَا أَضَاءَ لَهُمْ مَشَوْا فِيهِ

وَإِذَا أَظْلَمَ عَلَيْهِمْ قَامُوا وَلَوْ شَاءَ اللَّهُ لَذَهَبَ بِسَمْعِهِمْ


وَأَبْصَارَهُمْ إِنْ شَاءَ اللَّهُ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ (20)

صدق الله العلي العظيم

(سورة البقرة)

إقرار المشرف

أشهد أن إعداد هذه الرسالة الموسومة بـ (تمنجة تقلبات الحمل الكهربائي والتنبؤ به باستخدام
انحدار المتجه الداعم مع الخوارزميات الفوضوية في المنطقة الجنوبية) قد جرى تحت إشرافي
في كلية الإدارة والاقتصاد - قسم الإحصاء - جامعة البصرة، وهي جزء من متطلبات نيل درجة
الماجستير في الإحصاء.

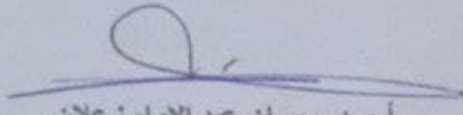

أم. ساهرة حسين زين النعبي

المشرف

2022/2/6

إقرار رئيس قسم الاحصاء

بناء على توصية المشرف أرشح هذه الرسالة للمناقشة.



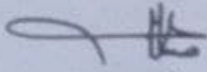
أ.م. د. رمضان عبد الامام زعلان

رئيس قسم الاحصاء

2022 / 2 / 6

إقرار الخبير اللغوي

أشهد أن الرسالة الموسومة بـ (تمنّجة نقلات الحمل الكهربائي والتنبؤ به باستخدام الحدار المتجه الداعم مع الخوارزميات الفوضوية في المنطقة الجنوبية) لطالبة الماجستير (هدى عبد السادة هاشم) قد تمت مراجعتها من الناحية اللغوية وأصبحت خالية من الأخطاء اللغوية والطباعية ولأجله وقعت.



أ.د. صباح عبد الكريم مهدي

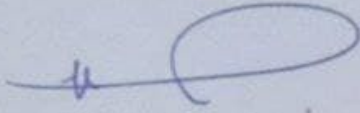
الخبير اللغوي

كلية الإدارة والاقتصاد - قسم الاقتصاد

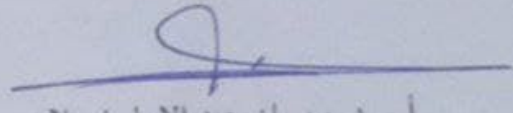
التاريخ ٢ / ١ / 2022

إقرار لجنة المناقشة

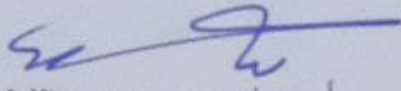
نشهد اننا أعضاء لجنة المناقشة قد اطلعنا على الرسالة الموسومة بـ (نمذجة تقلبات الحمل الكهربائي والتنبؤ به باستخدام انحدار المتجه الداعم مع الخوارزميات الفوضوية في المنطقة الجنوبية) وقد ناقشنا الطالبة (هدى عبد السادة هاشم) في محتوياتها وفيما له علاقة بها ونرى أنها جديرة بالقبول لنيل درجة ماجستير في الإحصاء وبتقدير (جيد جداً عالي).



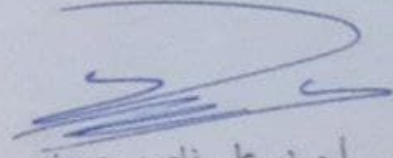
أ.م. د. مهند فائز كاظم
عضواً
2022/2/6



أ.م. د. ريسان عبد الامام زعلان
رئيس اللجنة
2022/2/6



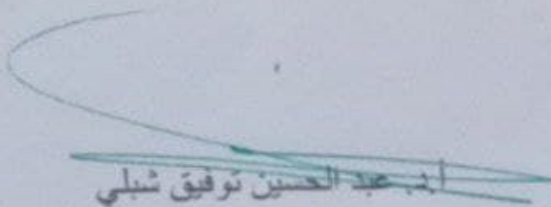
أ.م. ساهره حسين زين الثعلبي
عضواً مشرفاً
2022/2/6



أ.م. د. علي ناصر حسين
عضواً
2022/2/6

مصادقة مجلس الكلية

صدقت الرسالة بمجلس كلية الادارة والاقتصاد / جامعة البصرة



أ.م. د. عبد الحسين توفيق شبلي
عميد كلية الادارة والاقتصاد
التاريخ 2022/2/6

الإهداء

إلى بقية الله التي لا تخلو من العترة الهادية، إلى المعد لقطع دابر الظلمة، إلى المرتجى لإزالة الجور والعدوان، إلى باب الله الذي منه يُؤتى، إلى وجه الله الذي إليه يتوجه الأولياء، إلى السبب المتصل بين الأرض والسماء... إلى المضطر الذي يجب إذا دعا

"صاحب العصر والزمان الأمام المهدي (عجل الله فرجه)"

إلى من مهد طريق العلم بعد الله وذل لي الصعاب، فلم يبخل عليّ طيلة حياته، والذي لم يتهاون يوماً في توفير سبيل الخير والسعادة لي إلى ملاكي في الحياة

"أبي الموقر"

إلى من وقفت إلى جانبي وتحلت بالصبر والإيمان، إلى من أفضلها على نفسي، والتي كانت تعد الأيام بالدقائق لأنهي هذا الجهد المتواضع، رفيقة دربي وبلسم جراحي... إلى من ساندتني في صلاتها ودعائها

"أبي الحبيبة"

إلى من اعتمد عليه في كل كبيرة وصغيرة، سندي وعضدي، إلى من رافقني بمحبته وعطائه الحاني

"أخي مصطفى"

إلى من تمنيت وجودها معي في هذا اليوم، وفي كل يوم... إلى الانسانة العظيمة التي توسدت التراب

"روح جدتي العزيزة"

إلى من ضحوا بأنفسهم ودمائهم في سبيل ان ننام نوما هنيئاً، وأرواح من فقدنا منهم شهداء العراق

الباحثة ...

شُكْرٌ وَتَقْدِيرٌ

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

((رب اوزعني ان أشكر نعمتك التي أنعمت علي وعلى والدي وان أعمل صالحا ترضاه وأدخلني برحمتك في عبادك الصالحين))

الحمد لله والشكر له كما ينبغي لجلال وجهه وعظيم سلطانه، عدد خلقه ورضا نفسه وزنة عرشه ومداد كلماته على أن منَّ عليَّ بإنجاز هذه الدراسة، والصلاة والسلام على أفضل الخلق نبينا محمد "صلى الله عليه وآله وسلم" وعلى آله الطيبين الطاهرين وسلم تسليما كثيرا.

يطيب لي، بعد ان اتممت رسالتي المتواضعة، ان أتوجه بالشكر والثناء الى الأستاذ المساعد ومشرفتي على الدراسة " ساهره حسين زين الثعلبي" لتحملها عناء الاشراف على هذه الرسالة ومتابعتها المستمرة لي بالنصح والإرشاد والتوجيه، وما علمتني من فيض انسانيته وخلقه الرفيع، والتي كان لأشرافها العلمي الدقيق وتوجيهاتها السديدة وما جادت به نفسها الكريمة من خلق نير، الأثر البالغ في اظهار الرسالة وانجازها على نحو نافع ومفيد، وتبقى المفردة اللغوية قاصرة عن التعبير عما يختلج في داخلي من مشاعر الود والاحترام تجاهها، وفقها الله وجزاها الله خير الجزاء.

ويسرني ان أتقدم بالشكر الجزيل الى الأستاذ الفاضل الدكتور عبد الحسين توفيق شبلي عميد كلية الإدارة والاقتصاد المحترم والسيد رئيس قسم الإحصاء السابق الدكتور علي ناصر لما ابداه من جهود علمية في المرحلة التحضيرية والسيد رئيس القسم الحالي الدكتور ريسان عبد الامام زعلان لما منحه من وقت وجهد وتوجيه، عظيم الامتنان والدعاء للعلي القدير ان يجزيكما كل خير.

كذلك يحتم علي رد الجميل بالشكر والعرفان الى الست خديجة عدنان حميد المحترمة اذ امدتني بما احتجت إليه في بداية الدراسة.

كذلك أتقدم بالشكر والتقدير الخالص الى جميع اساتذتي الكرام في قسم الإحصاء وكل من ساهم في تعليمي.

كما أتقدم بالشكر والامتنان للأساتذة الافاضل رئيس وأعضاء لجنة المناقشين المحترمين على تفضلهم بقبول مناقشة هذه الرسالة وبما سوف يبدوه من ملاحظات واءاء علمية قيمة تسهم في اغناء هذه الرسالة. كما لا يسعني الا ان أتقدم بالشكر الجزيل لكل من مد لي يد العون، او اسدى لي معروفاء، او قدم لي نصيحة من قريب او بعيد ولو بالدعاء بظهر الغيب، او كانت له اسهامه صغيرة او كبيرة في انجاز هذه الدراسة، ومساندتهم لي طوال مدة الدراسة فلهم مني خالص الشكر والتقدير، وجزاهم الله عني الجزاء الأوفى.

وأتوجه بالشكر والتقدير لوالدي العزيزين اللذين غرسا فيّ حب العلم من الصغر، وقدما لي كل غالٍ ونفيس، وكان لهما الفضل بعد الله فيما وصلت الية الان فلا املك الا الدعاء لهما بطول العمر وحسن العمل وبلوغ الجنان.

فجزى الله الجميع عني خير جزاء المحسنين ووفقهم لما يحبه ويرضاه انه هو السميع العليم

ومن الله التوفيق والسداد

الباحثة ...

المحتويات

الصفحة	الموضوع
الفصل الأول الاطار المفاهيمي العام للدراسة	
1	1.1: المقدمة
3	1.1.1: مشكلة الدراسة
4	2.1.1: هدف الدراسة
4	3.1.1: الدراسات السابقة
الفصل الثاني: انحدار المتجه الداعم (Support Vector Regression)	
13	1.2: مفهوم انحدار المتجه الداعم
17	1.1.2: طريقة عمل انحدار المتجه الداعم
18	2.1.2: دالة النواة
20	3.1.2: خطوات انحدار المتجه الداعم
20	1.3.1.2- جمع مجموعه التدريب
21	2.3.1.2- اختيار نواة
21	3.3.1.2- بناء مصفوفة الارتباط
21	4.3.1.2- حل المصفوفة لمقدر النموذج
22	5.3.1.2- تشكيل المقدر
22	2.2: نظرية انحدار المتجه الداعم
26	1.2.2: مشكلة الانحدار
26	2.2.2: ميزات انحدار المتجه الداعم
27	3.2.2: نمذجة انحدار المتجه الداعم
27	4.2.2: مبدا تقليل المخاطر التجريبية
28	5.2.2: دالة الخسارة
29	6.2.2: أنواع انحدار المتجه الداعم

32	3.2: أساليب تعلم انحدار المتجه الداعم
32	1.3.2: المشكلة المزدوجة او مشكلة البرمجة التربيعية
34	2.3.2: تقدير أداء انحدار المتجه الداعم
36	3.3.2: تحديد معلمات انحدار المتجه الداعم
38	4.3.2: تحليل المعلمات في انحدار المتجه الداعم
39	5.3.2: خصائص انحدار المتجه الداعم
الفصل الثالث: نظرية الفوضى (Chaos theory) والخوارزميات الفوضوية الهجينة (Chaotic Hybrid algorithm)	
41	1.3: مفهوم الفوضى
42	1.1.3: نظرية الفوضى
45	2.1.3: عناصر نظرية الفوضى
45	1.2.1.3 - التنظيم الذاتي
45	2.2.1.3 - التشعبات
46	3.2.1.3 - الجاذب
47	3.1.3: الخصائص الأساسية للأنظمة الفوضوية
48	2.3: الخوارزميات الفوضوية الهجينة
48	1.2.3: الخوارزمية الجينية الفوضوية واستخدامها في تحديد معلمات نموذج انحدار المتجه الداعم
48	1.1.2.3: الخوارزمية الجينية
53	1.1.1.2.3: طريقة عمل الخوارزمية الجينية
52	1.1.1.1.2.3 - التهينة
52	2.1.1.1.2.3 - الانتقاء او الاختيار
54	3.1.1.1.2.3 - التقاطع
55	4.1.1.1.2.3 - الطفرة
55	5.1.1.1.2.3 - مقياس التوقف
56	2.1.1.2.3: دالة اللياقة البدنية
56	3.1.1.2.3: خصائص الخوارزمية الجينية
58	2.1.2.3: الخوارزمية الجينية الفوضوية

59	1.2.1.2.3: ميزات الخوارزمية الجينية الفوضوية
60	2.2.1.2.3: استخدام الخوارزمية الجينية الفوضوية في تحديد معالم نموذج انحدار المتجه الداعم
60	3.2.1.2.3: خطوات تنفيذ الخوارزمية الجينية الفوضوية المهجنة مع انحدار المتجه الداعم
63	4.2.1.2.3: تحسين معالم انحدار المتجه الداعم باستخدام الخوارزمية الجينية الفوضوية
63	2.2.3: خوارزمية المناعة الفوضوية واستخدامها في تحديد معالم انحدار المتجه الداعم
63	1.2.2.3: خوارزمية المناعة
64	1.1.2.2.3: خطوات عمل خوارزمية المناعة
68	2.1.2.2.3: ميزات خوارزمية المناعة
69	2.2.2.3: خوارزمية المناعة الفوضوية
70	1.2.2.2.3: خطوات عمل خوارزمية المناعة الفوضوية المهجنة مع انحدار المتجه الداعم
72	3.2.3: خوارزمية تحسين سرب الطيور واستخدامها في تحديد معالم نموذج انحدار المتجه الداعم
72	1.3.2.3: خوارزمية تحسين سرب الطيور
75	1.1.3.2.3: خطوات عمل خوارزمية تحسين سرب الطيور
77	2.1.3.2.3: معامل وزن القصور الذاتي
78	3.1.3.2.3: مزايا خوارزمية تحسين سرب الطيور
80	2.3.2.3: خوارزمية تحسين سرب الطيور الفوضوية
80	1.2.3.2.3: البحث المحلي الفوضوي
81	2.2.3.2.3: خطوات عملية خوارزمية تحسين سرب الطيور الفوضوية المهجنة في انحدار المتجه الداعم
83	3.2.3.2.3: تحديد دالة اللياقة البدنية
84	5.2.3.2.3: تحسين معالم انحدار المتجه الداعم استنادا الى خوارزمية تحسين سرب الطيور الفوضوية

85	3.3: النمذجة والتقلبات والتنبؤ
85	1.3.3: النمذجة
86	1.1.3.3: أهمية النمذجة
87	2.1.3.3: مراحل النمذجة
87	2.3.3: التقلبات في الحمل الكهربائي
89	3.3.3: مفهوم التنبؤ
90	1.3.3.3: خطوات بناء نموذج التنبؤ
91	2.3.3.3: أهمية التنبؤ
91	3.3.3.3: خصائص التنبؤ
92	4.3.3.3: مقاييس دقة التنبؤ
الفصل الرابع	
تقدير النموذج والتنبؤ باستهلاك الطاقة الكهربائية	
94	1.4: التحليل الوصفي للبيانات
94	1.1.4- وصف متغيرات الدراسة
95	2.1.4- الإحصاءات الوصفية
97	2.4 المفهوم العام للظاهرة المدروسة
99	3.4: نتائج التحليل الإحصائي
107	4.4: المفاضلة بين النماذج المستخدمة في الدراسة
109	5.4: نتائج التنبؤ
الاستنتاجات والتوصيات	
112	- الاستنتاجات
113	- التوصيات
مراجع الدراسة	
115	- المصادر

فهرس الجداول

الصفحة	العنوان	رقم الجدول
56	قيم العجلة المتدرجة	1
55	الطفرة	2
95	الإحصاءات الوصفية للمتغير المعتمد Y	3
96	الإحصاءات الوصفية للمتغيرات التوضيحية	4
100	تقدير معلمات (SVR) مع الخوارزمية الجينية الفوضوية (CGA)	5
101	تقدير معلمات (SVR) مع الخوارزمية الجينية الفوضوية (CGA) بعد عملية التكرار	6
103	تقدير معلمات انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA)	7
104	تقدير معلمات انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA) بعد عملية التكرار	8
106	تقدير معلمات انحدار المتجه الداعم (SVR) مع خوارزمية تحسين سرب الطيور الفوضوية (CPSO)	9
108	معايير المفاضلة بين النماذج	10
110	قيم التنبؤ المستقبلية باستهلاك الطاقة الكهربائية (2020 – 2028)	11

فهرس الاشكال والمخططات

الصفحة	العنوان	الرقم
14	تعيين مساحة الادخال x في مساحة ميزة عالية الابعاد	شكل (1)
16	عمل انحدار المتجه الداعم (SVR)	شكل (2)
17	تمييز SVM (يسار) و SVR (يمين)	شكل (3)
18	معلمت انحدار المتجه الداعم (SVR)	شكل (4)
18	تعيين مجموعات البيانات الى مساحة عالية الابعاد مع دالة النواه	شكل (5)
19	ثلاثة أنواع من دوال النواه (Kernal function)	شكل (6)
20	نموذج ومعالجة (SVR)	شكل (7)
20	استخدام النواة (Kernel)	شكل (8)
22	مساحة ميزة انحدار المتجه الداعم	شكل (9)
23	مفهوم انحدار المتجه الداعم (SVR)	شكل (10)
25	رسم تخطيطي لانحدار المتجه الداعم (SVR) باستخدام دالة الخسارة	شكل (11)
28	دوال الخسارة	شكل (12)
29	دالة الخسارة (خطي (يسار) وتربيعي (يمين))	شكل (13)
41	أسس نظام (Edward Loran)	شكل (14)
44	ديناميكيات الخريطة اللوجستية	شكل (15)
47	يمثل الأنواع الثلاثة من الجوانب	شكل (16)
54	انتقاء عجلة الروليت	شكل (17)
55	دورة الخوارزمية الجينية (GA)	شكل (18)
61	ترميز ثنائي للسلسلة (الكروموسوم)	شكل (19)
62	مثال مبسط لتمثيل المعلمت (C, ϵ, σ)	شكل (20)
64	مثال على القيم الاولية (الاجسام المضادة)	شكل (21)
66	عمليات التقاطع	شكل (22)
67	عمليات الطفرات	شكل (23)
77	حالة تحسين سرب الطيور (PSO) الاولية	الشكل (24)
102	نتائج رسم انحدار المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA)	شكل (25)

105	نتائج نموذج انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA)	شكل (26)
107	نتائج نموذج انحدار المتجه الداعم (SVR) مع خوارزمية تحسين سرب الطيور الفوضوية (CPSO)	شكل (27)
109	نتائج المفاضلة بين النماذج الثلاث	شكل (28)
111	رسم القيم التنبؤية لسلسلة بيانات الاستهلاك السنوي للطاقة الكهربائية (2020 – 2028)	شكل (29)
52	المخطط الانسيابي للخوارزمية الجينية (GA)	مخطط (1)
59	الشكل الانسيابي للخوارزمية الجينية الفوضوية (CGA)	مخطط (2)
68	الشكل انسيابي لخوارزمية المناعة (IA)	مخطط (3)
72	خوارزمية المناعة الفوضوية (CIA)	مخطط (4)
76	خوارزمية تحسين سرب الطيور	مخطط (5)
83	عمل خوارزمية تحسين سرب الطيور الفوضوية (CPSO)	مخطط (6)

ملخص الدراسة

يؤدي التنبؤ بالحمل الكهربائي دوراً رئيساً في إدارة نظام الطاقة والتخطيط لجدولة قدرة توليد الطاقة، ومن الممكن أن يوفر التنبؤ السنوي الدقيق بالحمل الكهربائي إرشادات جديرة بالثقة لتشغيل شبكات الطاقة والتخطيط، وللتنبؤ الدقيق بالحمل الكهربائي دور كبير وأكثر أهمية في إدارته استراتيجية أنظمة الطاقة الوطنية، نظراً لأن منحى الحمل الكهربائي غير خطي فإن التنبؤات السنوية للطاقة الكهربائية هي مشكلة غير خطية.

تطبق هذه الدراسة نموذجاً جديداً للتنبؤ بالحمل الكهربائي من خلال تطبيق انحدار المتجه الداعم (SVR) مع الخوارزمية الهجينة الفوضوية (Chaotic Hybrid algorithms) لتحسين أداء التنبؤ، والتي تعمل على حل مشكلات تحسين معاملات انحدار المتجه الداعم (SVR) كافة. إذ يعتمد استقرار نموذج (SVR) على اختيار المعلمات المثلى، فتم تهجين انحدار المتجه الداعم (SVR) مع ثلاث خوارزميات فوضوية (الخوارزمية الجينية الفوضوية (CGA)، خوارزمية المناعة الفوضوية (CIA)، خوارزمية تحسين سرب الطيور الفوضوية (CPSO)) لتحديد المعلمات المثلى والمناسبة لنموذج انحدار المتجه الداعم (SVR). ثم قورنت النماذج (SVRCGA)، (SVRCIA)، (SVRCPSO) بعضها مع بعض، لاختيار النموذج الأفضل واستخدامه في التنبؤ باستهلاك الطاقة الكهربائية للمدة (2020 – 2028) في المنطقة الجنوبية من العراق. وأظهرت النتائج أن نموذج انحدار المتجه الداعم مع الخوارزمية الجينية الفوضوية (SVRCGA) أكثر تفوقاً ودقة وكفاءة من نماذج التنبؤ الأخرى استناداً إلى معايير دقة التنبؤ الإحصائية.

الفصل الأول

الإطار المفاهيمي العام للدراسة

الفصل الأول: الإطار المفاهيمي العام للدراسة

1.1: المقدمة

تعد الطاقة الكهربائية من أهم مصادر الطاقة الرئيسية التي تسهم مساهمة كبيرة في جميع القطاعات التي تدفع بعملية التنمية إلى التقدم، فتعددت استخداماتها إذ دخلت في جميع مجالات الحياة، وتؤدي دوراً مهماً في عملية التنمية والرفاهية الاقتصادية والحضارية لاستخدامها كمستلزمات مهمة في الاقتصاد. لذا تعد أحد أهم مظاهر الحضارة والتطور ومقياساً للرفاهية في أي مجتمع من المجتمعات لما تؤديه هذه الطاقة من خدمات.

أن التنبؤ بالحمل الكهربائي ودقته قد حظي باهتمام متزايد، ويعد من المواضيع المهمة في علم الإحصاء نتيجة الحاجة إليه في مجالات الحياة المختلفة. كانت هذه العملية محط اهتمام الإحصائيين لفترة طويلة من الزمن لأن التنبؤ له أهمية واضحة ومؤثرة في دقة القرار، فمعظم الدول في برامجها التخطيطية تعتمد على أسس وطرائق علمية متقدمة للحصول على نتائج أكثر فعالية. اقترح العديد من الباحثين عدة أساليب للتنبؤ بالحمل الكهربائي لزيادة دقة الأداء وقد طوروا مجموعة متنوعة من الطرائق والخوارزميات الخاصة بعملية التنبؤ، من هذه الطرائق الأساليب التقليدية التي تضمنت نماذج المتوسطات المتحركة (ARIMA)، ونماذج (Box – Jenkins)، ونماذج التمهيد الآسي، والانحدار المتعدد، ونماذج تقدير بايزي (Bayesian)، واتسمت هذه الأساليب بضعف القدرة الملائمة غير الخطية، لذا من الصعب التنبؤ بالحمل الكهربائي باستخدام هذه الأساليب.

وكذلك الأساليب الذكية مثل تقنية الشبكة العصبية الاصطناعية (ANN)، كما تم مؤخراً تطوير خوارزمية انحدار المتجه الداعم (SVR) وهي آلة تعليمية جديدة قائمة على نظرية التعلم الإحصائي الذي يطبق مبدأ تقليل المخاطر الهيكلية (SRM) بدلاً من مبدأ تقليل المخاطر التجريبية (ERM) التي تستخدمه أغلب نماذج الشبكات العصبية الاصطناعية (ANN).

فتم استخدام انحدار المتجه الداعم (SVR) على نطاق واسع نظرا لأدائه المتميز غير الخطي لحل مشكلات التنبؤ في العديد من المجالات مثل التنبؤ بوصول السائحين والتنبؤ بالسلسلة المالية وغيرها، وقد اشارت استنتاجات هذه الأبحاث الى ان اختيار معاملات نموذج انحدار المتجه الداعم (SVR) تؤدي دورا في تحسين دقة التنبؤ بشكل كبير. ولإنشاء نموذج انحدار المتجه الداعم (SVR) يجب تعيين واختيار معلماته بعناية، ولتحقيق دقة التنبؤ بالأحمال الكهربائية فقد تم دمج مع الخوارزميات التطورية لتحديد المعلمات المناسبة لنموذج انحدار المتجه الداعم (SVR) لجعل النموذج الناتج أداة أكثر مثالية و كفاءة لزيادة الدقة التنبؤية بالحمل الكهربائي. الا انه لا يصل الى الامثلية، لوجود عيبين رئيسيين في هذه الخوارزميات إذ انها تتقارب ببطء، ويصل الحل الى المستوى الأمثل المحلي. لذلك، يتم التغلب على هذه العيوب عن طريق تهجين التسلسل الفوضوي مع الخوارزميات التطورية.

تقترح هذه الدراسة الية تعتمد على الخوارزمية الفوضوية لاختيار افضل القيم لمعاملات انحدار المتجه الداعم (SVR) والتي تمثل القيم المثالية للمعاملات في النماذج المستخدمة.

تم تقسيم محتوى الدراسة الى أربعة فصول وعلى وفق الآتي:

الفصل الأول: يعرض هذا الفصل الإطار المفاهيمي العام للدراسة واشتمل على ثلاثة مباحث، خصص المبحث الأول للمقدمة العامة، وضم المبحث الثاني الطاقة الكهربائية واهميتها في المنطقة الجنوبية، وضم المبحث الثالث النمذجة والتقلبات والتنبؤ.

الفصل الثاني: تطرق الفصل الثاني لانحدار المتجه الداعم واحتوى على ثلاثة مباحث، خصص المبحث الأول لمفهوم انحدار المتجه الداعم وطريقة عمله، وتناول المبحث الثاني نظرية انحدار المتجه الداعم، اما المبحث الثالث فقد ضم أساليب تعلم انحدار المتجه الداعم.

الفصل الثالث: عرض هذا الفصل نظريه الفوضى والخوارزميات الفوضوية الهجينة فقد خصص المبحث الأول لمفهوم الفوضى، واحتوى المبحث الثاني على الخوارزميات الفوضوية الهجينة.

الفصل الرابع: تناول هذا الفصل الجانب التطبيقي من الدراسة وشمل ثلاثة مباحث خصص المبحث الأول بالتحليل الوصفي للبيانات، وعرض المبحث الثاني تحليل السلسلة الزمنية، أما المبحث الثالث فقد شمل نتائج التحليل الاحصائي والتنبؤ بالحمل الكهربائي، واختتمت الدراسة ببعض الاستنتاجات والتوصيات.

1.1.1: مشكلة الدراسة

يعود سبب انخفاض الطاقة الكهربائية في العراق الى الحروب التي مرت بها البلاد وما تبعها من اعمال تخريب واضرار تعرضت لها محطات توليد الطاقة الكهربائية فضلا عن تآكل محطات توليد الطاقة الكهربائية لنقادمها. كل هذه الأسباب أدت الى الانقطاع المتكرر للتيار الكهربائي والذي يؤثر سلبا في معظم الأنشطة اليومية والاقتصادية للفرد، وأصبح التنبؤ بالحمل الكهربائي القضية الأكثر أهمية في إدارة الطاقة لبلد يعاني من عجز كبير في تجهيز الطاقة الكهربائية. وبالنظر لضعف الأساليب التقليدية للتنبؤ بالحمل الكهربائي والأداء غير المتناسق للشبكة العصبية الاصطناعية (ANN) عند استخدامها تبرز مشكلة هذه الدراسة في التنبؤ بالحمل الكهربائي، لذا تم استخدام انحدار المتجه الداعم (SVR) لحل مشكلات التنبؤ، إذ ان اختيار المعلمات (C, σ, ϵ) في نموذج انحدار المتجه الداعم (SVR) يؤثر بشكل كبير في دقة التنبؤ، حيث إن دقة التنبؤ تعاني من الافتقار الى المعرفة باختيار ثلاث معلمات (C, σ, ϵ) في نموذج انحدار المتجه الداعم (SVR). ويمكن أن تؤدي إعدادات المعلمات المختلفة الى اختلافات كبيرة في الأداء. لذلك، يعد اختيار المعلمات المثلى خطوة مهمة في تصميم (SVR). ومن ثمّ تستخدم الخوارزمية الفوضوية لتحديد قيم المعلمات المناسبة والمثالية لنموذج انحدار المتجه الداعم (SVR). لذا يتم تطبيق نموذج انحدار المتجه الداعم (SVR) الهجين مع الخوارزمية الفوضوية بنجاح في التنبؤ بالحمل الكهربائي للحصول على نتائج اكثر مثالية ودقة وكفاءة في التنبؤ بالحمل الكهربائي.

2.1.1: هدف الدراسة

الهدف من هذه الدراسة تكوين نماذج هجينة من انحدار المتجه الداعم (SVR) مع الخوارزميات الفوضوية (الخوارزمية الجينية الفوضوية، خوارزمية سرب الطيور الفوضوية، خوارزمية المناعة الفوضوية)، ثم المقارنة بين هذه النماذج لاختيار افضلها وفق معايير إحصائية، واستخدام النموذج المختار في التنبؤ باستهلاك الطاقة الكهربائية في المنطقة الجنوبية للمدة (2020 – 2028).

3.1.1: الدراسات السابقة:

توجد عدة دراسات تناولت المواضيع المتبعة في الأسلوب الاحصائي المستخدم في الدراسة الحالية للتنبؤ بالحمل الكهربائي ندرج بعضها على وفق الآتي:

- تناولت دراسة (Seiji kawauchi, 2004) [81] طريقة تنبؤ قصير المدى تعتمد على نظرية الفوضى لطلب على الطاقة، يتم استغلال إمكانات طريقة إعادة الاعمار المحلية بشكل كامل لتحقيق التنبؤ الدقيق، ولحل التدهور المتوقع بسبب الاحمال المتغيرة، تم دمج العديد من طرائق تصحيح الأخطاء لتحسين طريقة التنبؤ المقترحة واشتملت العينة على بيانات الحمل الكهربائي الحقيقي لعام واحد (2004) في اليابان، توصلت نتائج الدراسة الى انه بالإمكان تحقيق التنبؤ الدقيق والمستقر على مدار العام عند استخدام بيانات السلاسل الزمنية فقط لمدة 12 أسبوعا ماضيا.

-اقتרכת دراسة (Zhengjun Liu, 2005) [93] نموذجا للتنبؤ بأسعار الكهرباء على أساس نظرية الفوضى، إذ تم اقتراح نموذج التنبؤ بأثر الكهرباء العالمية والمحلية على أساس RNN وتتبع الاتجاه المتطور لنماذج التنبؤ المقترحة المجاورة، اشتملت العينة على بيانات سلسلة زمنية لأسعار الكهرباء في سوق نيوانجلاند وأوضحت نتائج الدراسة نجاح التنبؤ بسعر الكهرباء في سوق نيوانجلاند، مما يجعل بالإمكان تطبيق طريقة التنبؤ بأسعار الكهرباء المقترحة في سوق الطاقة الصيني للتنبؤ بسعر الكهرباء.

- اقترحت دراسة (Shu Fan, 2006) [58] خوارزمية من خطوتين تعتمد على شبكة (SOM) وانحدار المتجه الداعم (SVR) للتنبؤ بحمل الذروة لنظام الطاقة، تم تصنيف جميع عينات التدريب وتعيينها الى شبكة (SOM) باستخدام أسلوب تعلم غير خاضع للأشراف، ثم استخدم انحدار المتجه الداعم (SVR) لحمل الذروة لليوم التالي لملاءمة بيانات التدريب لكل مجموعة فرعية في المرحلة الثانية، وقد تميز هذا النموذج بالقدرة على التعامل مع عدم الاستقرار في السلسلة الزمنية للحمل الكهربائي كما انه يتسم بقوة فائقة ويمكن تعديله بسهولة تبعا لأنظمة الطاقة المختلفة. وشملت العينة بيانات حمل الذروة الكهربائية اليومي في مدينة نيويورك وبيانات الطقس في (Central park)، واختير شهران نموذجيان، يتوافق الشهر الأول مع شهر كانون الثاني (يناير) لعام 2004، وهو شهر شتاء يزداد فيه الطلب. كما يتوافق الشهر الثاني مع شهر يوليو لعام 2004، وهو شهر صيفي يرتفع فيه الطلب، وتمثلت البيانات المستخدمة للتنبؤ بحمل الذروة للشهر الأول (1 يناير 2003 الى 31 ديسمبر 2003)، والشهر الثاني (1 يوليو 2003 الى 30 يونيو 2004)، وأظهرت نتائج الدراسة فعالية وكفاءة النموذج المقترح مقارنة بالطرائق الأخرى.

- اقترحت دراسة (Herui Cui, 2008) [51] نموذجا جديدا للتنبؤ بأسعار الكهرباء على أساس نظرية الفوضى، لتحليل السلاسل الزمنية الفوضوية لالتقاط خصائص سلوك السعر المعقد، ثم التحقق من السمة الفوضوية لطلب الحمل الكهربائي، واشتملت العينة على البيانات التاريخية لأسعار الكهرباء لمنطقة ما في شمال الصين للمدة من 1 فبراير الى 31 مايو في عام (2007). أشارت النتائج إلى أن استخدام نظرية الفوضى في التنبؤ بسعر الكهرباء له تأثير جيد في حل مشكلة تعقيد أسعار الكهرباء ومشكلات الأنظمة الديناميكية غير الخطية، والتحكم في امدادات الطاقة الأمنة لمحطة الطاقة.

- قدمت دراسة (Yaozhou Yu, 2008) [76] طريقة جديدة للتنبؤ بالحمل الكهربائي باستخدام انحدار المتجه الداعم (SVR) وخوارزمية التجميع ل FCM، إذ تم اختيار عينات التدريب عن

طريق التجميع العام على وفق درجة التشابه بين عينات المدخلات مع مراعاة الخاصية الدورية لتغيير الحمل، شملت عينة الدراسة بيانات الحمل التاريخية والبيانات المنطقية للنيازك في تيانجين (TIANJIN)، الصين من 23 أكتوبر الى 3 نوفمبر لعام 2000، توصلت الدراسة الى ان دقة وسرعة التنبؤ بالحمل الكهربائي باستخدام آله المتجه الداعم (SVM) اعلى من انحدار المتجه الداعم (SVR) الفردي وطريقة (ANN)، وبالإمكان امتلاكها قابلية كبيرة للتطبيق.

- هدفت دراسة (Panayiotis G. Curtis, 2009) [62] إلى التنبؤ بالسلاسل الزمنية لمعدل الايجار لناقلة Aframax الحديثة التي تبلغ 105000 طن من الوزن الساكن في الساعة باستخدام اساليب التحليل غير الخطي ونظرية الفوضى، طبق تحليل السلاسل الزمنية غير الخطية لمؤشر سعر الايجار لناقله Aframax في اليونان للمدة من 31 مارس لعام 2000 حتى 30 نوفمبر لعام 2007 على أساس اسبوعي، توصلت نتائج الدراسة الى ان مجموعة البيانات للنموذج المقترح تتميز بأداء جيد للغاية في الكشف عن الخصائص الفوضوية لسلسلة مؤشر الوقت المستأجر لناقلة Aframax نظراً لأن أداء التنبؤ يعتمد على درجة الطبيعة الفوضوية لهذه الأنظمة .

- تقترح دراسة (Guifen Gong, 2010) [49] نهجاً جديداً يعرف باسم خوارزمية تحسين سرب الطيور الفوضوية (chaotic particle swarm optimization algorithm) مع انحدار المتجه الداعم (SVR) للتنبؤ بالعوائد المالية. فشملت العينة بيانات العوائد المالية في الفترة من 5 يناير 2003 الى 28 ديسمبر 2005 في بورصة شنغهاي، واطهرت نتائج الدراسة أن النموذج المقترح يفوق النهج الأخرى في التنبؤ بالعائدات المالية.

- هدفت دراسة (Wang Li-ying, 2010) [117] الى التنبؤ بالتفريغ اليومي باستخدام نموذج انحدار المتجه الداعم (SVR) ومقارنة نتائجه بطرائق التنبؤ الأخرى باستخدام بيانات التفريغ اليومية الحقيقية. واشتملت العينة على بيانات التفريغ اليومية لحوض نهر اليانغتسى في الصين

- للسنوات (1992-2000)، واستنتجت الدراسة أن التنبؤ باستخدام (SVR) يمكن أن يقلل بشكل كبير من أخطاء المتوسط النسبي والأخطاء التربيعية لمتوسط الجذر من التفريغ اليومي المتوقع.
- اقترحت دراسة (Bijaya Ketan Panigrahi, 2011) [48] نموذج (SVRCGA) للتنبؤ بالطلب على الهاتف وتقييم جدوى (CGA) في تحديد المعلمات لتحقيق تحسين دقة التنبؤ من خلال دمجها في نموذج انحدار المتجه الداعم (SVR)، وأشارت إلى فائده التسلسل الفوضوي لتحسين دقة التنبؤ من خلال تجنب الوقوع في الامثلية المحلية كما في نموذج (SVRGA)، فشملت عينة الدراسة بيانات الطلب على الهواتف النقالة الشهرية التاريخية (2006-2008)، وتم التوصل الى أن النموذج المقترح يتفوق على النماذج الثلاثة الأخرى الموجودة، وهي نموذج المتوسط المتحرك المتكامل التلقائي (ARIMA) ونموذج الانحدار العام للشبكات العصبية (GRNN) ونموذج (SVRGASA).
- هدفت دراسة (Shijie Ye, 2011) [133] الى التنبؤ بالحمل الكهربائي على المدى الطويل (LTIF) باستخدام انحدار المتجه الداعم (SVR)، وتقديم توصيات بشأن نتائج الحمل الكهربائي في المستقبل وواردات وصادرات الاحمال، كما يمكن لـ SVR اكتشاف الرابط بين الحمل الكهربائي والنتاج المحلي الإجمالي (GDP) من خلال ترجمة العلاقة غير الخطية في مساحة الادخال الاصلية ذات الابعاد المنخفضة الى مساحة ميزة ذات ابعاد اعلى لتحسين دقة التنبؤ، واشتملت عينه التنبؤ على البيانات السنوية للحمل الكهربائي والواردات والصادرات للأحمال والنتاج المحلي الإجمالي للمدة (1995-2008) في الصين كعينات لتعلم انحدار المتجه الداعم (SVR)، أظهرت نتائج التنبؤ التي تم الحصول عليها من انحدار المتجه الداعم (SVR) انها أكثر موثوقية وجدوى للدور الحالي الذي تؤديه الصين في التجارة الدولية في الركود الاقتصادي العالمي.
- هدفت دراسة (Yucheng Dong, 2011) [69] الى تقديم نموذج (SSVRCIA) انحدار المتجه الداعم (SVR) الذي يجمع بين آلية التكيف الموسمية وخوارزمية مناعية فوضوية (CIA)

لإيجاد التنبؤ بالأحمال الكهربائية الشهرية. ونظرا لعدم التمكن من الحفاظ على التنوع السكاني لمجموعة سكانية اولية تحت ضغط انتقائي لذا يمكن لهذه الخوارزمية (IA) البحث عن حلول في المساحة الضيقة ويكون الحل بعيدا عن الحل الأمثل العالمي (تقارب السابق لأوانه). تم استخدام خوارزمية المناعة الفوضوية المقترحة (CIA) على أساس خوارزمية تحسين الفوضى و(IA)، للتغلب على المشكلة المحلية المثلى السابق لأوانه في تحديد ثلاث معلمات لنموذج انحدار المتجه الداعم (SVR). شملت العينة بيانات الحمل الكهربائي الشهرية التاريخية من شمال شرق الصين وتوصلت الدراسة الى ان النموذج المقترح حقق نتائج تنبؤ أكثر دقة من نماذج (ARIMA) و (TF-ε-SVR-SA) لذا فهو يعد بديلا واعدة للتنبؤ بالحمولة الكهربائية.

- قامت دراسة (Yanfang Deng, 2012) [54] بتطبيق تقنية تعلم إحصائية لانحدار المتجه الداعم (SVR) مع خوارزمية جينية فوضوية هجينة (CGA) للتنبؤ بتدفق حركة المرور على المدى القصير في المناطق الحضرية، واستخدمت الخوارزمية المقترحة للتغلب على الأمثلية المحلية السابقة لأوانها في تحديد ثلاث معلمات لنموذج انحدار المتجه الداعم (SVR) وتمت مقارنة الأداء التنبؤي مع نماذج أخرى، وشملت العينة بيانات تدفق حركة المرور في الصين فوق نهر Wuhan Yangtze، واستخدمت (CGA) مشغل رسم الخرائط الفوضوية والقدرة البحثية لـ (GAs) لتحديد مجموعة المعلمات المناسبة في نموذج (SVR). أظهرت نتائج الدراسة أن النموذج المقترح صالح للتطبيق في ديناميكية قصيرة الأجل للتنبؤ بتدفق حركة المرور.

- هدفت دراسة (Fuh – Yuh Ju, 2013) [78] الى تطبيق نموذج تنبؤ بالكهرباء قائم على انحدار المتجه الداعم (SVR) مع خوارزمية هجينة هي خوارزمية البحث عن الجاذبية الفوضوية (CGSA) إذ تم تهجين البحث المحلي الفوضوي مع خوارزمية البحث عن الجاذبية في نموذج (SVR)، أي نموذج (SVRCGSA)، ثم دمجت الآلية الموسمية في نموذج (SVRCGSA)، مكونة نموذج (SSVRCGSA) لتحسين أداء التنبؤ الشهري للطلب على الكهرباء. اذ تستخدم

خوارزمية البحث الفوضوية (CGSA) المقترحة في البحث من خلال دالة رسم الخرائط الفوضوية اللوجستية في تكرار (GSA) الأصلي للبحث عن أفضل الحلول الحالية وتحسينها، فضلا عن تطبيق آلية موسمية للتعامل مع اتجاه الكهرباء الموسمية، ولإثبات دوال التهجين الفوضوي لرسم الخرائط من خلال استخدام خوارزميات التحسين في نموذج انحدار المتجه الداعم لما يتمتع به من قدره على تجنب التقارب قبل الأوان، شملت العينة بيانات الحمل الكهربائي الشهرية التاريخية لشمال شرق الصين، توصلت نتائج الدراسة الى أن النموذج المقترح يحقق نتائج تتبؤ أكثر دقة من طراز ARIMA ونماذج (TF - ε - SVR - SA) والإشارة الى ان عملية تهجين (CLS) والجمع بين الآلية الموسمية مفيدة لتحسين دقة التنبؤ.

- قدمت دراسة (Wei-Chiang Hong, 2013) [70] نموذجا للتنبؤ بالحمل الكهربائي قائم على انحدار المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA) لتحسين أداء التنبؤ. ويتم استخدام (CGA) المقترح استنادا الى خوارزمية تحسين الفوضى والخوارزمية الجينية (GA)، اذ تستخدم عشوائية داخلية من تكرارات الفوضى للتغلب على الامثلية المحلية السابقة لأوانها في تحديد ثلاث معلمات لنموذج (SVR)، اشتملت العينة بيانات الحمل الكهربائي الشهرية التاريخية لشمال شرق الصين، وشارت الدراسة الى ان التهجين بين التعديل الموسمي والتسلسل الفوضوي مفيد لتحسين دقة التنبؤ. وظهرت نتائج التنبؤ ان نموذج (SSVR CGA) اكثر دقة من نماذج (ARIMA) و (Tf - ε - SVR - SA)، لذا فهو بديل واعد للتنبؤ بالحمل الكهربائي.

- قامت دراسة (Malek Sarhani, 2015) [108] باستخدام الأسلوب الهجين المقترح (SVR-FS-PSO) للتنبؤ بالحمولة الكهربائية وإجراء اختيار الميزة من اجل تحسين أداء ELF وتحسين معلمات انحدار المتجه الداعم (SVR)، واشتملت العينة على بيانات الاحمال الكهربائية التاريخية المستخدمة في مسابقة (THE EUNITE) للمدة (1 كانون الثاني لعام 1997 الى 31 كانون الأول 1998) في المغرب، واستنتجت الدراسة ان الاسلوب الهجين المقترح لاختيار النموذج

يحقق قدرة تعميم أفضل بالمقارنة مع نموذج (SVR) الكلاسيكي في أثناء استخدام اختيار الميزة وبدون استخدامه. وان اختيار الميزة الأكثر صلة واختيار انحدار المتجه الداعم (SVR) بواسطة (PSO) يمكن ان يعزز دقة ELF، خاصة في حالة مجموعات البيانات الكبيرة التي تحتوي على العديد من الميزات او المتغيرات.

- تمت دراسة (Bing – Yi Lin, 2017) [86] لتطبيق دالة الفوضى ومفاهيم الحوسبة الكمومية (CQ) لمعالجة التراجعات المضمنة بما في ذلك عمليات الطفرات والتبادل والتحول للخوارزميات الجينية. واقترحت هذه الدراسة نموذجًا جديدًا للتنبؤ بحمولة الكهرباء عن طريق تهجين دالة الفوضى والحوسبة الكمومية (CQ) مع الخوارزمية الجينية (GA) في نموذج انحدار المتجه الداعم (SVR) المسمى (SVRCQGA) لتحقيق مستويات دقة تنبؤ أكثر. شملت العينة بيانات تحميل الكهرباء في الساعة من مسابقة التنبؤ العالمي للطاقة في مدينه تاويوان (Taoyuan City) لعام (2014)، اثبتت نتائج الدراسة أن نموذج (SVRCQGA) المقترح متفوق على نماذج تنافسية أخرى.

- اقترح (Ming – Wei Li, 2017) [88] نهجًا جديدًا وهو نموذج (SVRCQBA) للتنبؤ بالحمل الكهربائي من خلال تهجين نموذج انحدار المتجه الداعم (SVR) مع آلية الحوسبة الكمومية (CQ)، وخوارزمية الخفافيش المقترحة (BA)، للحصول على دقة تنبؤ أعلى اذ تستخدم دالة رسم الخرائط من Cat لتجنب التقارب السابق لأوانه في اثناء قيام الخوارزمية (QBA) بالنمذجة والمساهمة في أداء التنبؤ الدقيق. إذ شملت العينة بيانات التحميل الكهربائي في مسابقة التنبؤ بالطاقة في الصين (Jiangsu) لعام (2014) وتوصلت هذه الدراسة الى أن النموذج (SVRCQBA) المقترح يتفوق على النماذج البديلة الأخرى من حيث دقة التنبؤ، كما يسهم تهجين تقنيات التحسين الجديدة المختلفة بشكل أكبر للتنبؤ باستخدام خوارزميات تطويرية جديدة بتحسينات

الدقة إذ سيكون الاتجاه البحثي أكثر أهمية في عمل التنبؤ بالحمل الكهربائي المستند إلى انحدار المتجه الداعم (SVR) في المستقبل.

- قام (Yangquan Dong, 2018) [56] بدراسة تهجين نموذج انحدار المتجه الداعم (SVR) مع خوارزمية بحث الوقواق الفوضوي (SVRCCS) استنادًا إلى استخدام دالة رسم خرائط فوضوية لإثراء مساحة البحث عن الوقواق وتنويع السكان لتجنب الامتلية المحلية، وللتعامل مع الطبيعة الدورية للأحمال الكهربائية يتم دمج آلية موسمية مع نموذج (SVRCCS) لإعطاء انحدار المتجه الداعم (SVR) الموسمية مع نموذج بحث الوقواق الفوضوي (SSVRCCS) للحصول على أداء تنبؤ أكثر دقة. إذ تمتلك خوارزمية بحث الوقواق (CS) القدرة على المساهمة في تحقيق نتائج أكثر إرضاء في التنبؤ بالحمل الكهربائي. ومع ذلك، فإن خوارزمية (CS) الأصلية تعاني من عيوبها الكامنة فقد تتطلب المعلومات ضبطًا دقيقًا وفقدان التنوع السكاني، لذا اقترحت هذه الدراسة بعض آليات التحسين المهمة لتحديد تركيبات المعلمة المناسبة لنموذج انحدار المتجه الداعم (SVR). شملت العينة بيانات الحمل الكهربائي لمدة نصف ساعة من سوق كوينزلاند الإقليمية لسوق الكهرباء الوطنية (NEM، كوينزلاند، أستراليا). توصلت نتائج الدراسة إلى أن نموذج (SSVRCCS) المقترح يتفوق على النماذج البديلة الأخرى.

- قدمت دراسة (Peyman Yousef, 2020) [126] مقارنة جديدة لتحسين دقة النماذج في التنبؤ بقيم الاستهلاك اليومي للمياه في المناطق الحضرية مستندة إلى نظرية الفوضى للمساعدة في تحسين إدارة النظام، وخفض التكاليف وتحسين الموثوقية. اشتملت العينة على مجموعة بيانات من الاستهلاك الحضري لمدينة كيلونا في كندا كحالة اختبارية للتنبؤ بقيم الاستهلاك في المستقبل باستخدام فترات زمنية متفاوتة تحت جداول زمنية مختلفة لتحديد النماذج التي قد تُحسن أداء التنبؤ. حاولت هذه الدراسة إجراء دراسات حول تطبيق نظرية الفوضى في التنبؤ باستهلاك المياه،

وتوصلت النتائج الى تحسين أداء التنبؤ وتقديم ادلة لدعم مزيد من التحقيق في السلوك الفوضوي لقيم استهلاك المياه على مدد زمنية مختلفة.

تميزت الدراسة الحالية عن الدراسات السابقة في تصميم ثلاث نماذج هجينة ناتجة من دمج انحدار المتجه الداعم (SVR) مع الخوارزميات الفوضوية (الخوارزمية الجينية الفوضوية، خوارزمية تحسين سرب الطيور الفوضوية، خوارزمية المناعة الفوضوية) لتحديد المعلمات المثلى والمقارنة بين هذه النماذج واختيار أفضل نموذج وفق اختبارات احصائية ومن ثم التنبؤ به.

الفصل الثاني

انحدار المتجه الداعم

1.2: مفهوم انحدار المتجه الداعم

2.2: نظرية انحدار المتجه الداعم

3.2: أساليب تعلم انحدار المتجه الداعم

الفصل الثاني: انحدار المتجه الداعم (Support Vector Regression)

تمهيد:

تعد تقنية آلة المتجه الداعم (SVM) من التقنيات الحديثة والأكثر قوة وكفاءة لتحليل البيانات من أجل تصنيفها احصائياً وعمل تحليل الانحدار اللازم لها، تم تقديم (SVM) لأول مرة عام 1975 قبل أكثر من عقدين، ومنذ ذلك الحين فقد تمت دراستها بشكل مكثف إذ اكتسب جاذبية العديد من الباحثين والعلماء، وفي عام 1990 بدأ عصر البحوث النظرية المعتمدة على التعليم الآلي بسبب الخصائص المطلوبة التي تظهر مثل قدرات التعميم الجيدة والتقارب المضمون، والأساس النظري السليم، والتعامل الفعال مع البيانات عالية الأبعاد.

وتنقسم تقنية أو أسلوب المتجه الداعم (Support Vector Machine) إلى قسمين:

1. آلة المتجه الداعم للتصنيف (support vector machine for classification)

2. آلة المتجه الداعم للانحدار (support vector machine for regression) [74] وهو ما

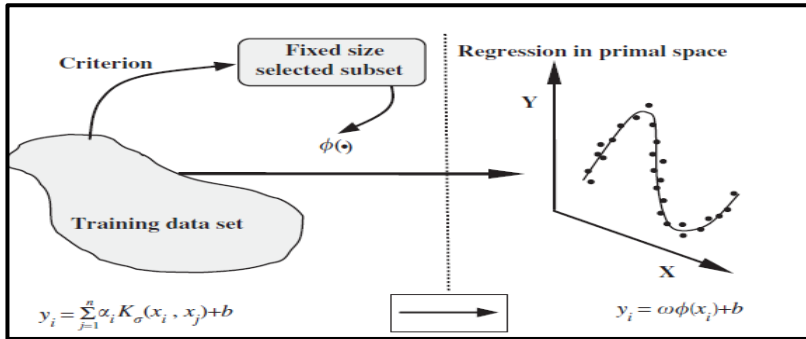
سيتم تناوله في هذه الدراسة.

1.2: مفهوم انحدار المتجه الداعم (Support Vector Regression)

تم تطوير آلة المتجه الداعم من قبل الباحث (Vapnik) في عام 1995 إذ ظهرت نسخة الانحدار من آلة المتجه الداعم (SVM) كأسلوب بديل وقوي لحل مشكلات الانحدار من خلال إدخال دالة خسارة بديلة، ويشار إلى هذا الإصدار باسم انحدار المتجه الداعم (SVR).

انحدار المتجه الداعم (SVR) نموذج مشتق من آلة المتجه الداعم (SVM)، فمهمة الانحدار هي مشابهة تماماً لمهمة التصنيف. وإن الهدف من نموذج آلة المتجه الداعم (SVM) هو إدارة مستوى بحيث تكون متجهات الدعم من مجموعتي التصنيف أبعد عن مستوى التصنيف، أما الهدف من نموذج انحدار المتجه الداعم (SVR) فهو العثور على مستوى الانحدار بحيث تكون جميع البيانات أقرب إلى (Hyperplane) [96].

تتبع صياغة انحدار المتجه الداعم (SVR) مبدأ تقليل المخاطر الهيكلية الذي يتمثل في تقليل الحد الأعلى لخطا التعميم بدلا من خطأ التنبؤ في مجموعة التدريب (مبدأ تقليل المخاطر التجريبية)، يتم ذلك عن طريق ادخال قيود على هيكل او انحاء مجموعة من الدوال التي يتم اجراء التقدير عليها. يزود انحدار المتجه الداعم (SVR) بإمكانية أكبر لتعميم علاقة المدخلات والمخرجات التي تم تعلمها من خلال مرحلة التدريب لعمل تنبؤات جيدة لبيانات الادخال الجديدة. يقوم انحدار المتجه الداعم (SVR) بتعيين بيانات الادخال x في مساحة مميزة عالية الابعاد F عن طريق التعيين غير الخطي لإنتاج وحل مشكلة الانحدار الخطي في مساحة المميزة



الشكل (1) تعيين مساحة الادخال x في مساحة مميزة عالية الابعاد. [46]

تقدير دالة الانحدار بناء على مجموعه بيانات معينة

$$G = (x_i, d_i)_i^n \dots (2.1)$$

حيث ان:

x : تشير الى متجه الادخال; d : قيمة المخرجات (الهدف); n : تمثل عدد عينات التدريب.

الهدف من النمذجة هو تحديد دالة الانحدار $y = f(x)$ التي تتنبأ بدقة بالمخرجات $\{d_i\}$ المقابلة

لمجموعة جديدة من امثلة المدخلات والمخرجات $\{(x_i, d_i)\}$ ويتم تقريب دالة الانحدار الخطي

باستخدام الدالة التالية:

$$f(x) = \omega \phi(x) + b \dots (2.2)$$

$$\phi: R^n \rightarrow F, \omega \in F$$

حيث ان:

ω, b : معاملات دالة الانحدار الخطي; $\phi(x)$: مساحه المميزة عالية الابعاد التي يتم تعيينها بشكل

غير خطي من مساحة الادخال x .

لذا فان الانحدار الخطي في مساحة الميزة عالية الابعاد (أي عدد كبير من الميزات) يستجيب للانحدار غير الخطي في مساحة الادخال ذات البعد المنخفض، متجاهلاً حساب المنتج الداخلي بين $\phi(x)$ و ω في مساحة الميزة عالية الابعاد، مما يمكن تقدير المعلمات b و ω عن طريق تقليل دالة المخاطر المنتظمة.

$$R_{SVR}(C) = R_{emp} + \frac{1}{2} \|\omega\|^2 = C \frac{1}{n} \sum_{i=1}^n L_S(d_i, y_i) + \frac{1}{2} \|\omega\|^2 \dots (2.3)$$

$$L_S(d, y) = \begin{cases} |d - y| - \varepsilon & ; \quad |d - y| \geq \varepsilon, \\ 0 & otherwise \end{cases} \dots (2.4)$$

حيث ان:

R_{SVR} : تمثل مخاطر الانحدار; R_{emp} : المخاطر التجريبية; C : تشير الى دالة تكلفة قياس

المخاطر التجريبية.

المصطلح الأول $\frac{\|\omega\|^2}{2}$ تدل على المسافة الإقليدية (هي الجذر التربيعي لمجموع الاختلافات المربعة بين العناصر المتجهة لمتغيرين).

نلاحظ في دالة المخاطر المنتظمة في معادلة (2.3)، مخاطر الانحدار (خطا مجموعة الاختبار)

R_{SVR} : هو الخطأ المحتمل الذي ترتكبه الدالة f في التنبؤ بالإخراج المقابل لمثال جديد (اختبار)

لمتجه الادخال.

$C(1/n) \sum_{i=1}^n L_S(d_i, y_i)$: يمثل الخطأ التجريبي والذي يسمى خطأ مجموعة التدريب الذي

تقدره دالة الخسارة غير الحساسة (ε) في معادلة (2.4).

المصطلح الثاني $\frac{\|\omega\|^2}{2}$ يمثل مصطلح التنظيم.

ويحسب الثابت C من خلال تحديد المقايضة بين الخطر التجريبي ومصطلح التنظيم، إذ يمثل

القدرة على التنبؤ بالانحدار، عند رفع قيمة C يزيد من أهمية المخاطر التجريبية بالنسبة لمصطلح

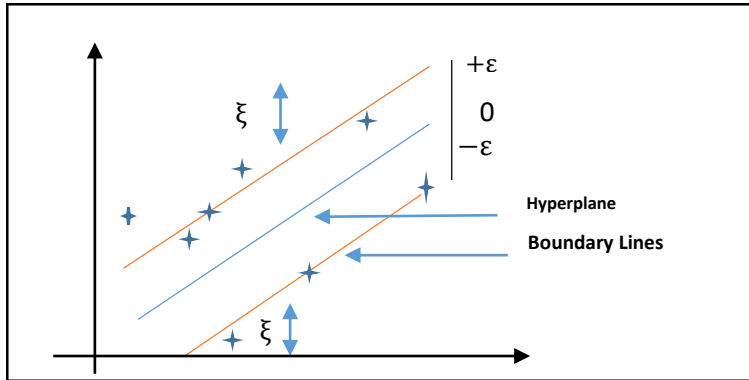
التنظيم [45]. فانحدار المتجه الداعم (SVR) كما يوحي من الاسم هو خوارزمية انحدار تدعم كلا

من الانحدار الخطي والانحدار غير الخطي، وهذا النهج يعمل على نظرية "انحدار المتجه الداعم

(SVR). يختلف هذا النهج عن آلة المتجه الداعم (SVM) بحيث تعد الأخيرة مصنفاً يستخدم

للتنبؤ بالتسميات القوية المنفصلة اما انحدار المتجه الداعم (SVR) فهو انحدار يتم استخدامه

للتنبؤ بالمتغيرات ذات الترتيب المستمر. والشكل (2) يوضح عمل (SVR)



الشكل (2) عمل انحدار المتجه الداعم (SVR). [95]

حيث ان:

Boundary Lines: تمثل الخطين المرسمين حول المستوي على مسافة ϵ والذي يستخدم لإنشاء

هامش بين نقاط البيانات.

(Hyperplane): عبارته عن خط يفصل بين فئتين من البيانات في بعد اعلى من البعد الفعلي.

ويتم تعريفه في انحدار المتجه الداعم (SVR) على انه الخط الذي يساعد في تنبؤ القيمة

المستهدفة والموضح في شكل (2).

مساحة الميزة (Feature Space): تشير مساحة الميزة الى مجموعات الميزات التي يتم

استخدامها لتوصيف البيانات الخاصة بك، وفي المتجه الداعم (SVM) تشير مساحة الميزة الى

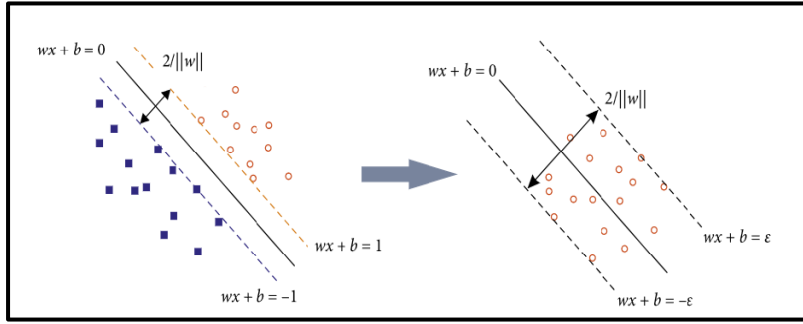
مجموعة مختلفة من الخصائص لوصف البيانات.

يعمل (SVR) على مبدأ آلة المتجه الداعم (SVM) مع اختلافات قليلة. إذ يحاول انحدار المتجه

الداعم (SVR) العثور على منحنى يستخدم لإيجاد التطابق بين المتجه وموضع المنحنى. ومن ثم

فان متجهات الدعم تساعد في تحديد أقرب تطابق بين نقاط البيانات والدالة التي يتم استخدامها

للتمثيل.



الشكل (3) تمييز SVM (يسار) و SVR (يمين). [95]

الفكرة الرئيسية لانحدار المتجه الداعم هو تقليل حدود خطأ التعميم والتي تشير إلى الخطر الحقيقي من إجراء التعلم الآلي. فقد طور Vapnik عام 1995 انحدار المتجه الداعم (SVR) من خلال دمج دالة الخسارة (ϵ - Insensitive) في المتجه الداعم (SVM) لحل مشكلات تقدير الانحدار غير الخطي أي أن الفكرة الأساسية لانحدار هي العثور على دالة يقع انحرافها عن البيانات الفعلية في المقياس المحدد مسبقاً، وتكون صيغته الهيكلية على وفق الاتي:

$$R(w) \leq Re mp(w) + \emptyset(n/h) \dots (2.5)$$

حيث أن:

$R(w)$: تشير الى ان المخاطر المتوقعة، يتم حسابها عبر مجموعة من الدوال $f(x, w)$.

$Re mp(w)$: يشير إلى المخاطر التجريبية في تصنيف العينات.

$\emptyset(n/h)$: تعني المخاطر السرية التي تمثل مستوى ثقة المصنف. [122]

1.1.2: طريقة عمل انحدار المتجه الداعم (SVR)

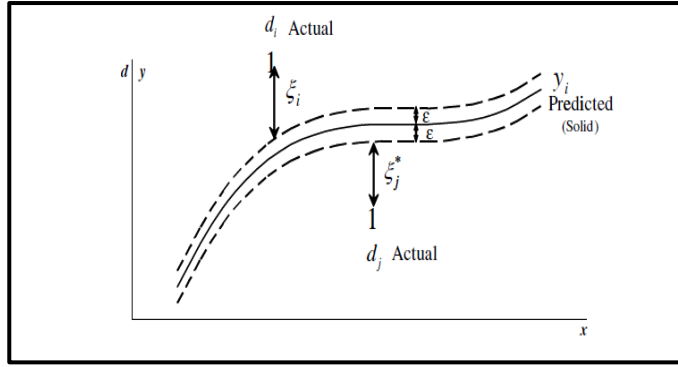
تكون طريقة عمل انحدار المتجه الداعم (SVR) على وفق الاتي:

$$y = f(x) = \langle w, x \rangle + b = \sum_{j=1}^M w_j x_j + b, y, b \in R, x, w \in R^M \dots (2.6)$$

$$f(x) = \begin{bmatrix} w \\ b \end{bmatrix}^T \begin{bmatrix} x \\ 1 \end{bmatrix} = w^T + b \quad x, w \in R^{M+1} \dots (2.7)$$

حيث ان: [95]

w^T : هو متجه الوزن ; y_i : يمثل قيمه الإخراج المرتبطة بـ x_i ، وكما هو مبين في شكل (4)



الشكل (4) معلمات انحدار المتجه الداعم (SVR). [102].

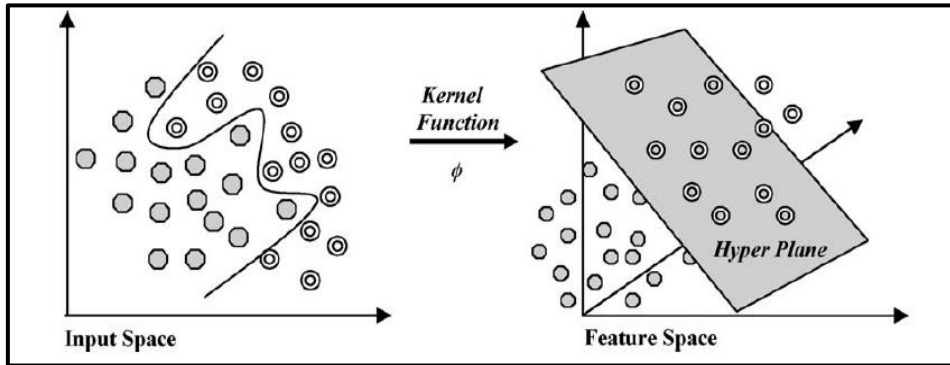
2.1.2: دالة النواة (kernel function)

يتم اجراء انحدار المتجه الداعم عند مساحة ذات بعد أعلى لذا يجب تكوين دالة وذلك بتحديد نقاط البيانات في ابعادها الأعلى وهذه الدالة تسمى النواة.

فالنواة (kernel) تعني تمكين اجراء العمليات التي يتعين القيام بها في مساحة المدخلات بدلا من

مساحة الميزة ذات الأبعاد العالية المحتملة لذا فالمنتج الداخلي لا يحتاج إلى تقييم في فضاء

الميزة. إذ تؤدي دالة النواة دورا مهما في المتجه الداعم وادائها. [75].



الشكل (5) تعيين مجموعات البيانات الى مساحة عالية الابعاد مع دالة النواة (Kernel). [80].

أذ توجد أربعة أنواع من دوال النواة المستخدمة في انحدار المتجه الداعم (SVR) وهي: خطي، متعدد الحدود، دالة الأساس الجذري (RBF)، ودالة السيني.

ويتم عرض النموذج الرياضي للدوال على النحو التالي:

- خطي (linear Kernel Function):

$$K_1(x, x_i) = x^T x_i \dots (2.8)$$

- نواة متعدد الحدود (polynomial kernel Function): وهي طريقة شائعة للنمذجة غير

الخطية وعادة ما تكون النواه الثانية هي الأفضل

$$K_2(x, x_i) = (\gamma x^T x_i + r)^d, \gamma > 0 \dots (2.9)$$

حيث ان:

d : هي درجة متعدد الحدود.

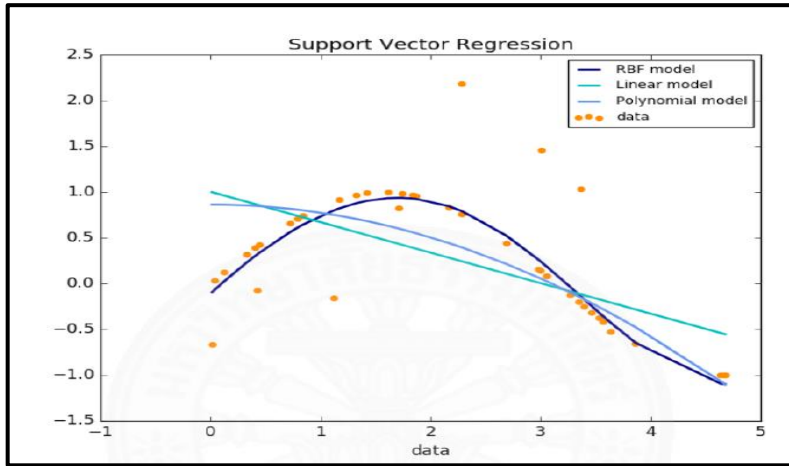
- دالة الأساس الجذري (Gaussian Radial Basis Function): وهي الدالة الأكثر شيوعاً مع شكل غاوسي (Gaussian) فضلا عن ذلك، فإن دالة نواة (RBF) ليست سهلة التنفيذ فحسب الا انها قادرة على تعيين البيانات بشكل غير خطي في مساحة الابعاد العالية ومن ثم فأنها تكون مناسبة للتعامل مع المشكلات غير الخطية.

$$K_3(x, x_i) e^{-\frac{\gamma \|x-x_i\|^2}{2\sigma^2}}, \gamma > 0 \dots (2.10)$$

حيث ان: ($\sigma > 0$) هو عرض النواه ; σ : هي معلمة يتم تحديدها يدوياً عادة.

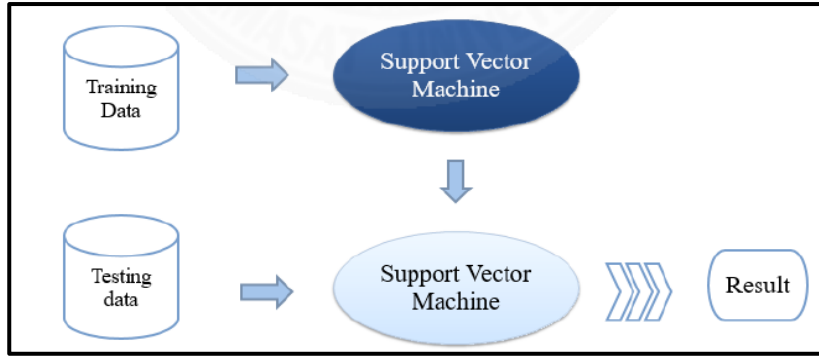
- داله النواه السيني (sigmoidal kernel Function): تنتج دالة الأساس حلا خطيا ومقطعاً كما يمكن ان يكون جذابا عندما تكون الانقطاعات مقبولة. [75]

$$K_4(x, x_i) = \tanh(\gamma x^T x_i + r) \dots (2.11)$$



شكل (6) ثلاثة أنواع من دوال النواه (Kernal function). [63].

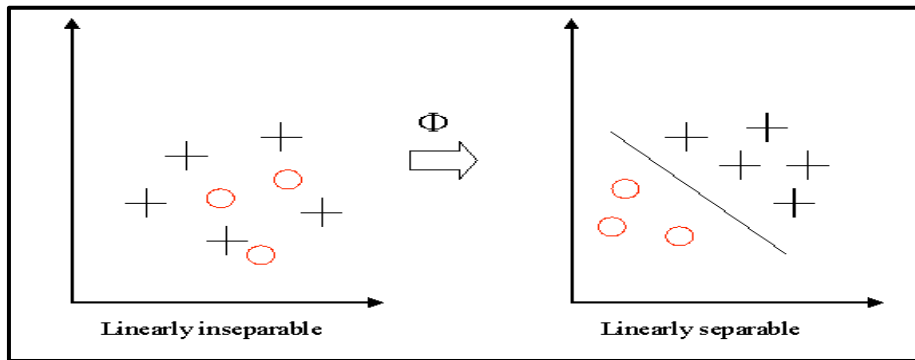
كما يوضح الشكل (7) نموذج ومعالجة انحدار المتجه الداعم (SVR)



شكل (7) نموذج ومعالجة (SVR). [63].

عندما تكون البيانات خطية، يتم تقسيم البيانات باستخدام مستوى فاصل (hyperplane) مختلف. الا انه تكون البيانات بعيدة كل البعد عن الخطية وان مجموعات البيانات لا يمكن فصلها. فتستخدم (Kernel) لتعيين خريطة غير خطية لبيانات الادخال إلى مساحة عالية الأبعاد،

عندئذ يكون التعيين الجديد مقسما خطيا. وكما موضح في شكل (8)



الشكل (8) استخدام النواة (Kernel). [75].

إذ يتم تعريف هذا التعيين من خلال النواة (Kernel) [75] [89]

$$K(x, y) = \phi(x) \cdot \phi(y) \dots (2.12)$$

3.1.2: خطوات انحدار المتجه الداعم (SVR)

توجد عدة خطوات لاستخدام انحدار المتجه الداعم يتم ادراجها على وفق الآتي: [95]

1.3.1.2- جمع مجموعه التدريب

تتألف مجموعة التدريب من العينات التي تم الحصول عليها لأجراء التنبؤ لها، إذ ان ميزات مجموعة التدريب يجب ان تكون مناسبة للمجال المتوقع. فعلى الرغم من ان انحدار المتجه الداعم (SVR) غير قادر الا على الاستيفاء بين البيانات الموجودة في مجموعة بيانات التدريب.

ولتوضيح ذلك، لنفترض ان انحدار المتجه الداعم (SVR) يستخدم لتقدير دالة جيب التمام (cosine) فان المجال الذي يجب التدريب عليه هو $(0 - 2\pi)$ ، فإذا تم تقييم التقدير خارج المجال الذي قمنا بتقييمه، فسنعتمد النتيجة مباشرة على اختيار النواة المستخدمة لانحدار المتجه الداعم (SVR) وإجراءات التحسين.

2.3.1.2- اختيار نواة (Kernel)

بناء على المشكلة يمكن اختيار أي نوع من أنواع النواه مثل (Sigmoid Kernel) و (Polynomial Kernel) و (Gaussian Kernel).

3.3.1.2- بناء مصفوفة الارتباط

من السهل تشكيل مصفوفة الارتباط، إذ يتم تقييم جميع أزواج النقاط في مجموعة التدريب وبعد التقييم تتم إضافة منظم النتائج في القطر، ليتم الحصول على المصفوفة وتكون على وفق الصيغة الاتية:

$$K_{i,j} = \exp\left(\sum_k \theta_k |x_k^i - x_k^j|^2\right) + \varepsilon \delta_{i,j} \dots (2.13)$$

حيث ان:

ε : يمثل المنظم ; $\delta_{i,j}$: تمثل دالة دلتا (Kroeniker).

فالمصفوفة تكون شبه محددة وسهلة الاستخدام، وتستخدم للتعبير عن مصفوفة الارتباط للبيانات في مساحة ابعاد أعلى من تلك التي تم اشتقاقها من مجموعة التدريب.

4.3.1.2- حل المصفوفة لمقدر النموذج

يعد الجزء الأساسي من تطوير انحدار المتجه الداعم (SVR)، ويمكن تعريفها على وفق الاتي:

$$\bar{K} \vec{\alpha} = \vec{Y} \dots (2.14)$$

$$\vec{\alpha} = \bar{K}^{-1} \vec{Y} \dots (2.15)$$

حيث ان:

\vec{Y} : تمثل قيم المتجهات التي تتوافق مع مجموعة التدريب ; \bar{K} : تمثل مصفوفة الارتباط.

$\vec{\alpha}$: تمثل مجموعه غير معروفة تحتاج الى حل ; \bar{K}^{-1} : تمثل معكوس مصفوفة الارتباط.

5.3.1.2 - تشكيل المقدر

عند معرفة قيمة $\vec{\alpha}$ يتم استخدام المقدر بسهولة. وحساب قيمة \vec{K} وتحديد المنتج الداخلي لها باستخدام متجهات الفا لتقدير y^* لنقاط x^* .

$$y^* = \vec{\alpha} \cdot \vec{k} \dots (2.16)$$

ولحساب عناصر مصفوفة k ، يتم استخدام الصيغة التالية:

$$k_i = \exp\left(\sum_k \theta_k |x_k^i - x_k^*|^2\right) \dots (2.17)$$

نستنتج من المعادلة (2.17) أنه إذا قمنا بعكس المصفوفة وقيمة المنظم (regularizer) صفراً،

فان المقدر يمر عبر جميع نقاط التدريب بشكل مثالي، لان \vec{k} يساوي الصف في k اذا كانت

نقطة الاختبار في مجموعة التدريب. لذا لا نحتاج الى استعادة حل مجموعة التدريب وبالإمكان

اغلاقها الى ان تتم بإضافة المنظم او الحصول على المعلمات المطلوبة.

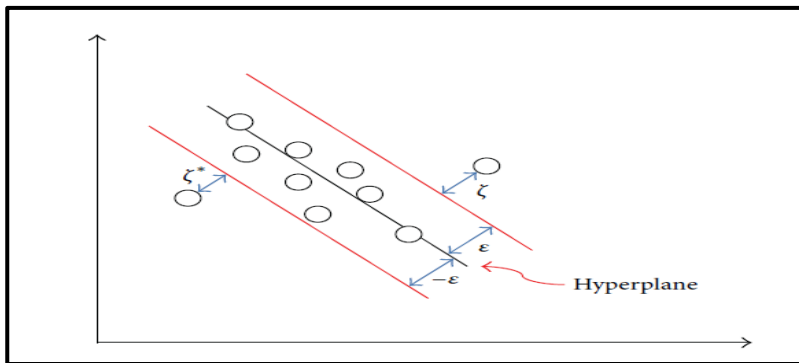
بعد اجراء الخطوات كافة، يكون نموذج انحدار المتجه الداعم (SVR) جاهزاً للتنبؤ. [95]

2.2: نظرية انحدار المتجه الداعم (Theory Support Vector Regression)

إن المفهوم الأساسي لانحدار المتجه الداعم (SVR) هو ان مجموعة البيانات الاصلية غير

الخطية (x_i) يتم تعيينها الى فضاء ميزه عالية الابعاد (أي عدد كبير من الميزات)، يوضح

الشكل (11) خريطة البيانات الى فضاء الميزة عالية الابعاد.



الشكل (9) مساحة ميزة انحدار المتجه الداعم (SVR). [45]

لتوضيح مفهوم (SVR) تمت صياغة مشكلة الانحدار على النحو التالي: نفرض ان مجموعة

بيانات تدريب معطاة

$$s = \{(x_1, y_1), \dots, (x_1, y_1)\} \dots (2.18)$$

حيث أن

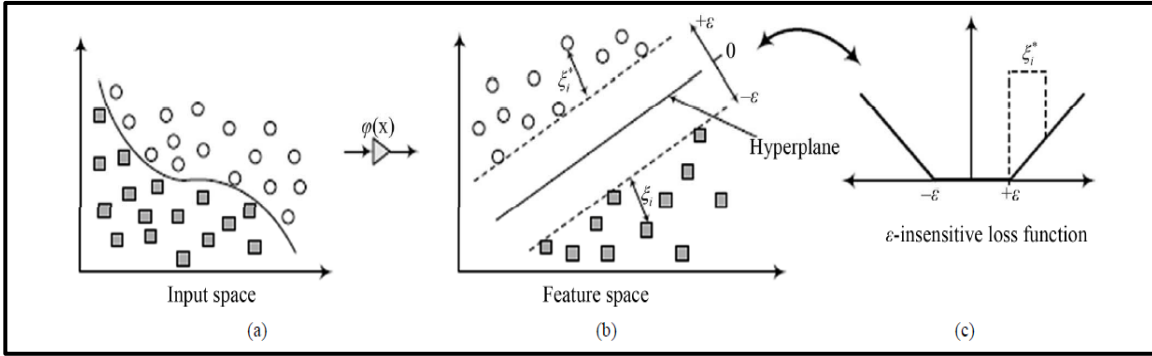
y_i : هو قيمة الإخراج المرتبطة بـ x_i .

وان دالة الانحدار (Support vector Regression) وهي

$$f(x) = w \times \Phi(x) + b \dots (2.19)$$

حيث أن:

$\Phi(x)$: يشير الى دالة التعيين غير الخطية ; b : يمثل مصطلح التحيز. [127]



(c) دالة الخسارة غير الحساسة، (b) التحويل الى دالة خطية، (a) مساحة الادخال غير الخطية.

الشكل (10) مفهوم انحدار المتجه الداعم (SVR). [133]

إن الهدف من (SVR) هو إيجاد دالة $f(x)$ التي تمتلك اقصى انحراف (ϵ) عن الأهداف (y_i) لجميع بيانات التدريب، إذ يتم تقديم دالة الخسارة لضمان تباين المتجه الداعم، وتعرف دالة الخسارة على وفق الاتي:

$$L_s(y_i, f(x)) = \begin{cases} |y_i - f(x_i)| - \epsilon, & |y_i - f(x_i)| \geq \epsilon \\ 0 & \text{otherwise} \end{cases} \dots (2.20)$$

حيث ان $L_s(y_i, f(x))$: تمثل دالة الخسارة التي يتم اختيارها بشكل مناسب، إذ تشير الى كيفية معاقبة الاختلافات بين y و $f(x)$ أي (القيم المتبقية).

فالخسارة تساوي صفراً إذا كان خطأ التنبؤ بالقيم اقل من ϵ ، فيما عدا ذلك فإنها تساوي قيمة اكبر

من ϵ ، كما هو الحال مع مشكلة التصنيف، يمكن إضافة متغيرات الركود غير سالبة (ξ_i, ξ_i^*) لتمثيل المسافة من القيم الحقيقية الى القيم المقابلة لـ ϵ لذا يمكن صياغة النموذج المقيد وعلى وفق الاتي:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i, \xi_i^*) \dots (2.21)$$

$$\begin{aligned} \text{subject to} \quad & y_i - [w \times \Phi(x)] - b \leq \varepsilon + \xi_i \\ & [w \times \Phi(x)] + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \quad ; \quad i = 1, 2, 3, \dots, l \end{aligned}$$

حيث أن:

$C \sum_{i=1}^m (\xi_i, \xi_i^*)$: يمثل المصطلح التجريبي الذي يقيس دالة الخسارة ε ، C : تمثل دالة تكلفه

قياس المخاطر التجريبية ξ_i, ξ_i^* : تمثل متغيرات الركود.

يتم حل مشكله التحسين المقيد باستخدام نموذج لاكرانج على وفق الاتي:

$$\begin{aligned} \text{Maximize} \quad & \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j \\ & - \alpha_j^*) K(x_i, x_j) \dots (2.22) \end{aligned}$$

$$\text{subject to} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad \alpha_i, \alpha_i^* \in [0, C] \quad , i = 1, 2, \dots, l$$

حيث أن:

α_i, α_i^* : تمثل مضاعفات لاكرانج $K(x_i, x_j)$: تمثل دوال النواه.

يتم تعريف دالة النواه (Kernal Function) على وفق الاتي:

$$K(x_i, x_j) = \Phi(x_i), \Phi(x_j) \dots (2.23)$$

باستخدام دوال النواه (Kernal) يمكن حساب انحدار المتجه الداعم (SVR) بدون تعيين صريح

في مساحة الميزة. [127]

ومن خلال حل المشكلة يمكننا الحصول على الحل الأمثل

$$\bar{\omega} = \sum_{i=1}^l (\bar{\alpha}_i^* - \bar{\alpha}_i) x_i \dots (2.24)$$

$$\bar{\alpha}^* = (\bar{\alpha}_1, \bar{\alpha}_1^*, \dots, \bar{\alpha}_l, \bar{\alpha}_l^*)^T \dots (2.25)$$

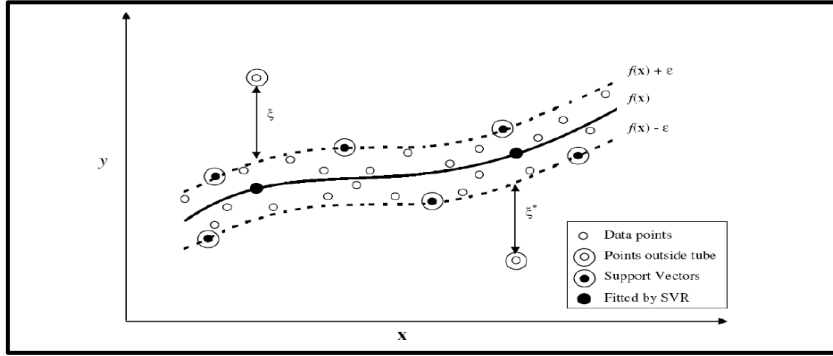
وتكتب دالة الانحدار الخطي على النحو الاتي:

$$f(x) = \sum_{i=1}^l (\alpha_i^* - \bar{\alpha}_i) K(x, x_i) + \bar{b} \dots (2.26)$$

إن الأسلوب المستخدم لحل المشكلة غير الخطية هو إسقاط مجموعة من البيانات إلى مساحة

الأبعاد العالية (العديد من الميزات) عن طريق إدخال دالة النواة. [120]

ميزه هذه الصيغة معادلة (2.23) و (2.26) ان شكل K يكون مفهوماً وبسيطاً للعديد من الاختيارات للمجموعة $\{\phi(x_i)\}$ ، بما في ذلك مجموعة الابعاد اللانهائية. على وفق ذلك، يمكن حساب المنتج النقطي في مساحة الميزة $(i \dots l)$ بدون تعيين المتجهين x_i و x_j في تلك المساحة (أي حساب $(\phi(x_i), \phi(x_j))$).



شكل (11) رسم تخطيطي لانحدار المتجه الداعم (SVR) باستخدام دالة الخسارة الحساسة [82]. يتم الحصول على معاملات α_i^* و α_i عن طريق حل مشكلة البرمجة التربيعية الآتية:

$$\text{Maximize: } R(\alpha^*, \alpha) = -0.5 \sum_{i,j=1}^p (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j) - \varepsilon \sum_{i=1}^p (\alpha_i^* + \alpha_i) + \sum_{i=1}^p (\alpha_i^* - \alpha_i) \dots (2.27)$$

والتي تخضع للقيود الآتية:

$$\sum_{i=1}^p (\alpha_i^* - \alpha_i) = 0 \quad , \quad 0 \leq \alpha_i, \alpha_i^* \leq C$$

بعد تقدير α^* و α و b باستخدام خوارزمية برمجة تربيعية مناسبة، تأخذ دالة الانحدار المستندة الى انحدار المتجه الداعم (SVR) الصيغة الآتية:

$$f(x, w) = f(x, \alpha, \alpha^*) = \sum_{i=1}^p (\alpha_i^* - \alpha_i)K(x_i, x) + b \dots (2.28)$$

تفسر المعلمات α و α^* بشكل بديهي وتمثل قوى تدفع وتسحب لتقدير انحدار $f(x)$ نحو القياسات y_i .

وتحسب معلمة التحيز b على وفق الآتي:

$$b = \begin{cases} y_i - f(x_i), & , b = 0 - \varepsilon \\ y_i - f(x_i), & b = 0 + \varepsilon \end{cases} \text{ , for } \alpha_i, \alpha_i^* \in (0, C) \dots (2.29)$$

حيث ان:

y_i : يمثل الناتج المستهدف المقابل.

C, ϵ : تمثل معلمتين تحددان من قبل المستخدم في صيغة انحدار المتجه الداعم (SVR)، إذ

تمثل C المفاضلة بين خطأ التقريب وتعقيد النموذج، أما ϵ فتشير إلى عرض المنطقة غير

الحساسة المستخدمة لتتناسب مع بيانات التدريب. [83]

1.2.2: مشكلة الانحدار (Regression Problem) [121]

لنفرض انه تم الحصول على مجموعه بيانات تدريب

$$s = \{(x_1, y_1), \dots, (x_N, y_N)\} \dots (2.30) \quad ; \quad x_n \in X \subseteq R, \quad y_n \in Y \subseteq R$$

مع N فيتم رسم عينات (x_n, y_n) بشكل مستقل وموزعة بشكل متماثل على وفق التوزيع

الاحتمالي غير المعروف ولكنه ثابت $P(x, y)$ ، وجعل مساحة الفرضية هي فئة الدوال الخطية في مساحة الابعاد (n) وكما في الصيغة الاتية:

$$H = \{h|h(x) = \langle w, x \rangle + b ; w, x \in R; b \in R\} \dots (2.31)$$

تستخدم مشكلة التعلم لتحديد الفرضية التي تقلل من المخاطر الهيكلية

$$R(h) = \int L(y, h(x)) dP(x, y) \dots (2.32)$$

استنادا الى مجموعه التدريب S فان $L(y, h(x))$ هي دالة الخسارة التي يتم اختيارها بشكل

مناسب، إذ تشير إلى كيفية معاقبة الاختلافات بين y و $h(x)$ أي (القيم المتبقية)، كما يمكن

محاولة اختيار فرضية تقلل من الحد الأعلى للمخاطر عن طريق المبدأ التعريفي (SRM).

2.2.2: ميزات انحدار المتجه الداعم (support vector regression) [85]

يتصف انحدار المتجه الداعم (SVR) بعدة مزايا من أهمها:

1. يكون انحدار المتجه الداعم (SVR) قوياً بالنسبة للقيم المتطرفة.
2. استخدام مصنفات متعددة تم تدريبها على أنواع مختلفة من البيانات وباستخدام قواعد الاحتمال.
3. تحديث نموذج القرار بسهولة.
4. يحسن من دقة التنبؤ عن طريق قياس الثقة في التصنيف.
5. تنفيذ طريقة انحدار المتجه الداعم (SVR) لكونها سهلة نوعاً ما.

6. يقوم انحدار المتجه الداعم (SVR) بإجراء حساب أقل مقارنة من تقنيات الانحدار الأخرى.

3.2.2: نمذجة انحدار المتجه الداعم (support vector regression)

(SVR) هو امتداد لنموذج تصنيف لـ "SVM" الذي يعتمد على نظرية التعلم الإحصائي الحديثة.

يستخدم مفهوم تقليل المخاطر الهيكلية (SRM) في صياغة انحدار المتجه الداعم (SVR) على

عكس منهجية تقليل المخاطر التجريبية إلى أدنى حد (ERM) الذي يستخدم نطاقاً واسعاً في

طرائق التعلم الآلي الإحصائي وتدريب (ANNS). يعمل مبدأ تقليل المخاطر التجريبية (ERM)

على تقليل خطأ التنبؤ في بيانات التدريب إلى الحد الأدنى على عكس مبدأ تقليل المخاطر الهيكلية

(SRM) يتم تقليل الحد الأعلى لخطأ التعميم. وهذا يعطي انحدار المتجه الداعم (SVR) قدرة

أكبر لتعميم العلاقة بين المدخلات والمخرجات التي تم تعلمها خلال مرحلة التدريب من أجل إجراء

تنبؤات دقيقة لبيانات الإدخال الجديدة.

يعد (SVR) أسلوباً خطياً في مساحة ميزة عالية الأبعاد (أي العديد من الميزات)، إذ ترتبط بشكل

غير خطي بمساحة الإدخال، في حين أن الخوارزمية الخطية تعمل في مساحة ذات بعد أعلى من

الميزات، إن استخدام النواة (Kernell) لا يتضمن أي حسابات في هذا الفضاء، لذا يتم إجراء

جميع الحسابات اللازمة مباشرة في مساحة الإدخال. [83]

4.2.2: مبدأ تقليل المخاطر التجريبية (ERM)

لا يمكن حساب المخاطر $R(h)$ لأن التوزيع الاحتمالي $p(x, y)$ غير معروف لخوارزمية التعلم

الإحصائية (statistical learning theory)، مع ذلك يمكن حساب تقدير يسمى بالمخاطر

التجريبية من خلال حساب متوسط دالة الخسارة في مجموعة التدريب.

$$p_{emp}(x, y) = \frac{1}{N} \sum_{n=1}^N \delta(X - X_n) \delta(Y - Y_n) \dots \quad (2.33)$$

حيث ان:

$p_{emp}(x, y)$: دالة الكثافة التجريبية ; N : عدد امثلة التدريب ; X_n : تمثل متجهات التدريب ;

Y_n : تمثل قيمة الهدف.

تكون دالة المخاطر التجريبية على وفق الاتي:

$$R_{emp}(h) \triangleq \int L(y, h(x))_{p_{emp}}(x, y) dx dy = \frac{1}{N} \sum_{n=1}^N L(y_n, h(x_n)) \dots (2.34)$$

حيث ان:

$R_{emp}(h)$: تمثل المخاطر التجريبية للدالة h .

والجدير بالذكر ان مبدا تقليل المخاطر التجريبية يشير الى ان الفرضية \hat{h} لتقليل المخاطر التجريبية يجب اختيارها من قبل خوارزمية التعلم الاحصائية

$$h_{ERM}^* = arg \min_{h \in H} R_{emp}(h) \dots (2.35)$$

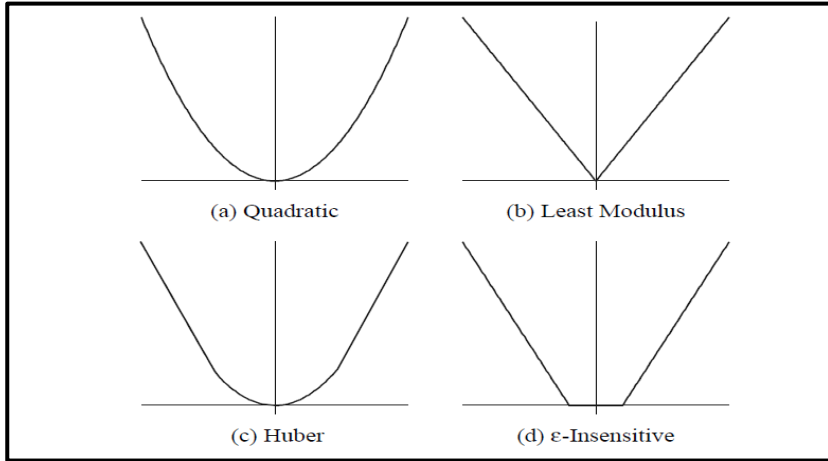
حيث ان:

H : تمثل قاعدة الدوال الواردة في فئة الفرضية. [121]

5.2.2: دالة الخسارة (The Loss Function)

بين الباحث Vapnik عام 1995 العديد من دوال الخسارة لانحدار المتجه الداعم (SVR)،

والشكل (12) يوضح أربعة أنواع من دوال الخسارة المحتملة



شكل (12) دوال الخسارة. [60]

في الشكل (9) (a) تشير الى دالة الخسارة على وفق معيار خطأ المربعات الصغرى التقليدية،

وفي الشكل (b) دالة خسارة لابلاسية (Laplacian loss function) أقل حساسية من دالة

الخسارة التربيعية للقيم المتطرفة، كما ان العالم Huber اقترح دالة خسارة قوية لها خصائص مثلى

عندما يكون التوزيع الأساسي للبيانات غير معروف والمتمثلة بالشكل (c). ان هذه الدوال الثلاث لن تنتج أي تناثر في متجهات الدعم، لذا اقترح Vapnik عام 1995 دالة الخسارة والمتمثلة في الشكل (d) كتقريب لدالة خسارة Huber التي تتيح الحصول على مجموعه متفرقة من متجهات الدعم. [60]

وتكون صيغة دالة الخسارة الخطية (ϵ - Insensitive) $L_1^\epsilon(y_i, f(x))$ على وفق الاتي:

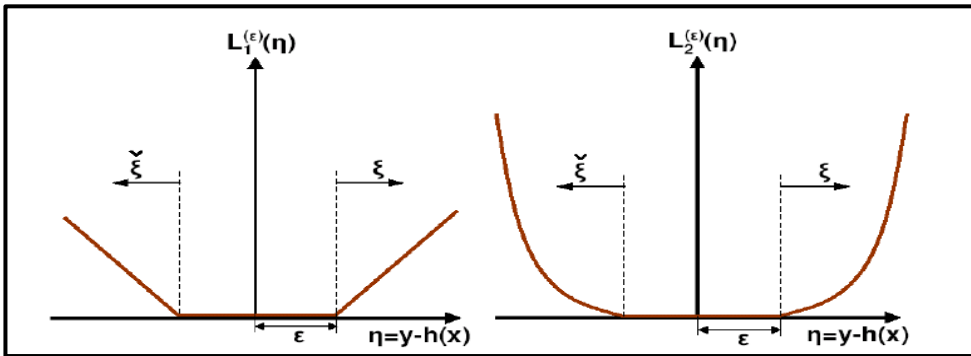
$$L_1^\epsilon(y_i, f(x)) = |y - f(x)|_\epsilon \triangleq \max(0, |y - f(x)| - \epsilon) \dots (2.36)$$

حيث ان:

h : تمثل دالة ذات قيمة حقيقية في المجال $x \in X$, و $y \in R$

وبالإمكان إعطاء دالة الخسارة من الدرجة الثانية على وفق الاتي:

$$L_2^{(\epsilon)}(y, f(x)) = |y - f(x)|_\epsilon^2 \triangleq \max(0, (y - f(x) - \epsilon)^2) \dots (2.37)$$



الشكل (13) دالة الخسارة (ϵ - Insensitive) خطي (يسار) وتربيعي (يمين). [121]

لذا لا يتم حساب عينة التدريب كخطأ طالما بقيت اقل من بعض القيمة المعطاة ϵ ، ويتم تحديد

المقدار الذي تفشل به عينة التدريب في الحصول على دقة الهدف ϵ من خلال متغيرات الركود

$$\xi((x_n, y_n), h, \epsilon) = \xi_n \triangleq \max(0, y_n - h(x_n) - \epsilon) \dots (2.38)$$

$$\check{\xi}((x_n, y_n), h, \epsilon) = \check{\xi}_n \triangleq \max(0, h(x_n) - y_n - \epsilon) \dots (2.39)$$

إذ ان ξ و $\check{\xi}$ تتوافق مع عينات التدريب التي تقع فوق الفرضية وتحتها. [121]

6.2.2: أنواع انحدار المتجه الداعم (SVR)

قبل التطرق الى أنواع انحدار المتجه الداعم (SVR) يفترض الإشارة الى ان هناك نوعين لآلة

المتجه الداعم (SVM) هما C-SVM و nu-SVM إذ يكون الفرق بينهما هو دالة الخسارة التي

يجب تحسينها وخاصة (Hyperparameter) الذي يحدد الخسارة المراد حسابها، ويتم توضيح

هذين النوعين على وفق الاتي:

- النوع الأول (C - svm): يتم تقليل دالة الخطأ اثناء التدريب لهذا النوع من (SVM)

$$\frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \dots (2.40)$$

$$\text{S.T: } y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, i = 1, \dots, N$$

حيث ان:

ξ_i : يمثل محددات لمعالجة البيانات (المدخلات) التي يصعب فصلها.

i : يمثل تسميات حالات تدريب N .

تستخدم النواة (ϕ) لتحويل البيانات من الادخال الى مساحة الميزة. إذ ان C هو عامل تحديد

الخطأ فكلما كانت قيمة C كبيرة فإن الخطأ (i) موجود، لذا يجب اختيار العامل C بدقة لتجنب

المبالغة في تقدير التكاليف.

- النوع الثاني (nu - svm): يقوم هذا النوع بتقليل دالة الخطأ [47]

$$\frac{1}{2} w^T w - \nu \rho + \frac{1}{N} \sum_{i=1}^N \xi_i \dots (2.41)$$

$$\text{S.T: } y_i(w^T \phi(x_i) + b) \geq \rho - \xi_i, \xi_i \geq 0, i = 1, \dots, N, \rho \geq 0$$

يكون انحدار المتجه الداعم (SVR) على أربعة أنواع وكالاتي:

- النوع الأول (epsilon - SVR): تمثل دالة الخطأ على وفق الاتي:

$$\frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \dots (2.42)$$

$$\text{S.T: } w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^*$$

$$y_i - w^T \phi(x_i) - b_i \leq \varepsilon + \xi_i$$

$$\xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, N$$

- النوع الثاني (nu - SVR): اقترح الباحث Scholkopf عام 1998 خوارزمية (nu - SVR)، إذ تم ضبط المعلمة (epsilon) تلقائياً، فتسمح هذه الخوارزمية بالتحكم في عدد متجهات الدعم عن طريق تحديد نسبة عدد متجهات الدعم الواردة في الحل الى العدد الإجمالي للعينات في مجموعة البيانات. [83]، ويمكن تمثيل دالة الخطأ على وفق الآتي: [47]

$$\frac{1}{2} w^T w - C \left(v\varepsilon + \frac{1}{N} \sum_{i=1}^N (\xi_i + \xi_i^*) \right) \dots (2.43)$$

$$\text{S.T: } (w^T \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i$$

$$y_i - (w^T \phi(x_i) + b_i) \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, N, \quad \varepsilon \geq 0$$

ويكمن الاختلاف عند التحكم فيه في أثناء عملية التدريب فنحتاج في حالة epsilon-SVR الى

التحكم في الحد الأقصى للخطأ المسموح به لإعداد الانحدار، ويكون هذا الامر جيداً في إيجاد

أفضل قالب ممكن دون الخوض في الكثير من العمليات الحسابية. فلا يمكن التحكم في عدد

متجهات الدعم المستخدمة في أثناء التحسين، لأنه احتمال أن تكون قليلة، او تكون كثيرة وفي هذه

الحالة قد تحتاج إلى الكثير من الموارد.

اما nu-SVR، فانه لا يمكن التحكم في الحد الأقصى لمقدار الخطأ المسموح به في النموذج لذا

ستتحكم في عدد متجهات الدعم، لذا نحتاج الى موارد ابسط واقل من epsilon-SVR لكن من

المحتمل أن تجد خطأ أكبر قليلاً اعتماداً على احتياجاتنا، لذلك يجب اختيار نوع من انحدار المتجه

الداعم (SVR) يناسب مشكلة التعلم الآلي والتي تحاول إيجاد حل أمثل.

- انحدار المتجه الداعم الخطي (Linear SVR)

تقوم دالة النواة (kernel function) بتحويل البيانات الى مساحة الميزة ذات ابعاد اعلى لجعل

الفصل الخطي ممكناً [47]

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot \langle x_i, x \rangle + b \dots (2.44)$$

- انحدار المتجه الداعم غير خطي (Non-Linear SVR)

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot \langle \varphi(x_i), \varphi(x) \rangle + b \dots (2.45)$$

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b \dots (2.46)$$

3.2: أساليب تعلم انحدار المتجه الداعم (Support Vector Regression)

تتشارك أساليب انحدار المتجه الداعم (SVR) بالمزايا والعيوب نفسها لالة المتجه الداعم (SVM)، أي ان طرائق تدريب آلة المتجه الداعم (SVM) يمكن تطبيقها على انحدار المتجه الداعم (SVR) عندما لا تعتمد بشكل مباشر على خصوصيات مشكلة (SVM)، ولا يمكن تمديد الا عدد قليل من أساليب (SVM) الى انحدار المتجه الداعم (SVR) لان معظمها يعتمد على الصيغة الرياضية لـ (SVM).

ذكر الباحث (Collobert et al) في عام 2001 نسخة من تقنية (SVM) لمشكلات (SVR)، إذ بدأ الباحثون بالمسألة الأساسية في معادلة (2.21) واستمروا في معالجة المشكلة الأساسية وإعادة صياغة النموذج المزدوج (2.27) باستخدام صيغة مصفوفة لمتجهات من اجل الحصول على مشكلة تقليل البرمجة التربيعية (QP) وعلى وفق الاتي:

$$\begin{aligned} \min_{\alpha^*, \alpha} Q(\alpha^*, \alpha) &= \frac{1}{2} (\alpha^* - \alpha)^T K (\alpha^* - \alpha) - (\alpha^* - \alpha)^T d + \epsilon (\alpha^* - \alpha)^T \mathbf{1} \dots (2.47) \\ \text{s. T} \quad (\alpha - \alpha^*)^T \mathbf{1} &= 0 \\ \alpha_i, \alpha_i^* &\leq C, -\alpha_i, -\alpha_i^* \leq 0 \\ \text{for} \quad i &= 1, 2, 3, \dots \dots N \end{aligned}$$

حيث ان:

K : تمثل مصفوفة النواه (kernel matrix). [103]

1.3.2: المشكلة المزدوجة او مشكلة البرمجة التربيعية (quadratic programs)

الفكرة الأساسية هي بناء دالة لاكرانج (Lagrange function) على وفق دالة الهدف والقيود المقابلة لها، من خلال توفير مجموعة من المتغيرات المزدوجة، وعلى وفق الاتي:

$$L = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) - \sum_{i=1}^l \alpha_i (\varepsilon + \xi_i - Y_i + \langle \omega, x_i \rangle + b) - \sum_{i=1}^l \alpha_i^* (\varepsilon + \xi_i^* + Y_i - \langle \omega, x_i \rangle - b) \dots (2.48)$$

حيث أن:

L : تمثل لاكرانج (Lagrange) ; $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ تمثل مضاعفات لاكرانج (Lagrange).

يجب ان تفي المتغيرات المزدوجة بالقيود الإيجابية، أي $\alpha_i^{(*)}, \eta_i^{(*)} \geq 0$

لاحظ ان $\alpha_i^{(*)}$ تشير الى α_i و α_i^* المشتقات الجزئية لدالة لاكرانج (L) يجب ان تتلاشى فيما يتعلق بالمتغيرات الأولية $(\omega, b, \xi_i, \xi_i^*)$ وذلك من اجل تحقيق المثالية

$$\partial_b L = \sum_{i=1}^l (\alpha_i^* - \alpha_i) = 0 \dots (2.49)$$

$$\partial_\omega L = \omega - \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i = 0 \dots (2.50)$$

$$\partial_{\xi_i^{(*)}} L = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \dots (2.51)$$

باستبدال $\partial_b L$ و $\partial_\omega L$ و $\partial_{\xi_i^{(*)}} L$ في معادلة (2.47) يعطي مشكلة التحسين المزدوج

$$\text{Maximize} \begin{cases} -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle \\ -\varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \end{cases} \dots (2.52)$$

$$\text{subject to} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad \text{and} \quad \alpha_i, \alpha_i^* \in [0, C]$$

من خلال الاشتقاق الجزئي لمشكلة التحسين المزدوج في (2.51) تتم إزالة المتغيرات المزدوجة

من خلال الشرط $\partial_{\xi_i^{(*)}} L = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0$ ، وتتم إعادة صياغتها على النحو

الآتي:

$$\eta_i^{(*)} = C - \alpha_i^{(*)} \dots (2.53)$$

يمكن إعادة صياغة معادلة (2.49) على وفق الآتي: [111]

$$\omega = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i \dots (2.54)$$

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \dots (2.55)$$

تتسم المشكلة المزدوجة بالمزايا الآتية: [100]

1. ان مشكلة التحسين هي مشكلة برمجة تربيعية مع القيود الخطية، وهي سهلة الحل وتضمن

حلا فريدا.

2. يظهر متجه الادخال فقط داخل المنتج النقطي، مما يضمن ان ابعاد مساحة الادخال يمكن أن

تكون مخفية عن الحسابات المتبقية. إذ يتحول فضاء الإدخال إلى فضاء ذي أبعاد عالية، فإن

الحساب لا يحدث في ذلك الفضاء بل في الفضاء الخطي.

3. يسمح النموذج المزدوج باستبدال المنتج النقطي لمتجهات الإدخال بتحويل غير خطي لمتجه

الإدخال.

2.3.2: تقدير أداء انحدار المتجه الداعم (SVR)

يعتمد أداء ودقة تقنية انحدار المتجه الداعم (SVR) على الاعدادات الجيدة للمعلمات

(معلمة $(C, \epsilon, \text{Kernel})$ ، مثل نوع النواة (Kernel) ونوع دالة الخسارة ومعلمات النواة (Kernel).

ونظرا لان تعقيد نموذج المتجه الداعم وأداء التعميم يعتمد على جميع المعلمات، عندئذ يصبح

اختيار المعلمة المثلى أكثر تعقيدا.

تستخدم المعلمة (ϵ) لتلائم بيانات التدريب ويمكن ان تؤثر قيمة هذه المعلمة في عدد متجهات

الدعم المستخدمة في التنبؤات فكلما زاد حجم المعلمة يقل عدد متجهات الدعم، مما يؤدي الى

تقديرات انحدار اقل تعقيدا، وتتحكم المعلمة ϵ في عرض المنطقة الحساسة، ومن جهة أخرى فان

قيمها الكبيرة تؤدي الى تقديرات ثابتة اكثر، فان قيم (ϵ, C) تؤثر في تعقيد النموذج ولكن بطريقة

مختلفة.

يعتمد تحديد نوع معاملات دالة (kernel) على التطبيق المراد القيام به إذ يجب ان يعكس توزيع قيم الادخال (x) لبيانات التدريب، وتحدد المعلمة C المفاضلة بين تعقيد النموذج ودرجة الانحرافات الأكبر من C في صياغة التحسين، مثلا إذا كانت قيمة C كبيرة جدا (ما لانهاية) فان الهدف يتلخص في الحد من المخاطر التجريبية فقط بدلا من التعقيد النموذجي في الصياغة

المثلى. [83]

يجب تحسين اختيار هذه المعلمات بشكل صحيح لتقليل خطأ التعميم، اقترح Cherkassky و Ma عام (2004) تقنية مختلفة تعتمد على بيانات التدريب، يتم حساب قيم C باستخدام بيانات التدريب بدلا من اللجوء الى اعاده اخذ العينات، وباستخدام طريقة التقدير الاتي:

$$C = \text{Max}(|\bar{Y} + 3\sigma_y|, |\bar{Y} - 3\sigma_y|) \dots (2.56)$$

حيث ان:

\bar{Y} : يمثل الوسط الحسابي ; σ_y : يمثل الانحراف المعياري لقيم Y لبيانات التدريب،

وهذا النهج لديه ميزه ان القيم المتطرفة المحتملة مستقرة.

اشار Cherkassky و Ma الى ان قيمة (ε) يجب ان تكون متناسبة مع الانحراف المعياري لضوضاء الادخال، باستخدام نظرية الحد المركزي، إذ ان (ε) تعطى على وفق الاتي:

$$\varepsilon = 3\sigma \sqrt{\frac{\ln n}{n}} \dots \dots (2.57)$$

حيث ان:

σ : الانحراف المعياري لضوضاء الادخال.

وبما ان قيمة σ غير معروفة، يتم تقديرها على وفق الاتي:

$$\hat{\sigma} = \sqrt{\frac{n^{1/5}K}{n^{1/5}K - 1} \cdot \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \dots (2.58) \quad ; \quad 2 \leq k \leq 6$$

حيث ان:

K: تمثل تقديرات التحيز المنخفض/التباين العالي ; \hat{y} : تمثل القيمة المتوقعة لـ y من خلال

ملاءمة الانحدار الخطي لبيانات التدريب لتقدير تباين الضوضاء (noise variance). [100]

3.3.2: تحديد معلمات انحدار المتجه الداعم (Support Vector Regression)

نظرا لخصائص انحدار المتجه الداعم (SVR) العديدة والأداء الواعد للتعميم، فقد حقق نجاحا كبيرا في كل من المنصات الأكاديمية والصناعية، ويتم ادراج بعض خصائصه الرئيسية وعلى وفق الآتي:

1. أهمية الحل باستخدام (SVR) فريدة ومثاليه ومستحيلة لإنتاج الحد الأدنى المحلي، لذا فان

طريقة تدريب (SVR) تشبه حل مسائل البرمجة التربيعية المقيدة خطيا.

2. يستخدم فقط نقاط البيانات المطلوبة لحل دالة الانحدار، مما يؤدي الى تذبذب الحل.

3. لدية القدرة على محاكاة نموذج العلاقات غير الخطية.

والجدير بالذكر أن يكون (SVR) محدودا في مجال البحوث الأكاديمية والتطبيقات الصناعية،

على الرغم من الميزات الفائقة التي يتمتع بها لان المستخدم يحدد المعلمات على النحو المناسب،

كما يتم تعيين معلمات انحدار المتجه الداعم بعناية لبناء نموذج كفوء. تؤدي المواصفات المختلفة

للمعلمات الى حدوث اختلافات كبيرة في الأداء، لذلك فان احدى الخطوات المهمة في بنية انحدار

المتجه الداعم (SVR) هي تحديد المعلمة المثلى. لذا يتوقف نجاح أداء (SVR) وكفاءته على

اختيار المعلمات المثلى الصحيحة (C, ϵ, σ) فتحديد المعلمات المثلى لمجموعة بيانات معينة هي

المسألة الرئيسية، ولا توجد قواعد عامة لتحديد هذه المعلمات. [46]

تتفاقم مشكلة اختيار المعلمات المثلى بسبب تعقد نموذج (SVR) ويتوقف ادائه العام على

المعلمات الثلاث معا، إذ ان تحسين المعلمة بشكل منفصل لا يكفي للحصول على نموذج انحدار

المتجه الداعم (SVR) وانما بتحسين تفاعلهم معا.

أشارت نتائج جميع الدراسات الى اختيار ثلاث معلمات في نموذج انحدار المتجه الداعم (SVR)

تؤدي دورا حاسما في تحسين دقة التنبؤ، إذ تعاني دقة التنبؤ الضعيفة من الافتقار الى المعرفة

باختيار هذه المعلمات (C, ϵ, σ) ، وان استقرار أي نموذج من نماذج انحدار المتجه الداعم (SVR) ودقته يتوقف على اختيار المعلمات الأساسية إذ لا يمكن بناء نموذج عالي الفعالية الا بعد اختيار هذه المعلمات بعناية. [42]

من الممكن تلخيص نهج العملية القائمة لاختيار المعلمات المثلى على وفق الآتي:

- من اجل العثور على معلمات انحدار المتجه الداعم (SVR) المثلى اقترح Smola في عام 2002 طريقة تحسين البحث في الشبكة، الا ان هذه الطريقة تستغرق وقتا طويلا.
- المعلمتان (C, ϵ) اختارهما العديد من الباحثين (Cherkassky & Mulier, 1998) (Burges, & Smola, 1999; Vapnik, 1995) على أساس التجربة والمهارات السابقة، ولا تتناسب هذه الأساليب بالتأكيد مع المستخدمين غير الخبراء.
- اقترح (Mattera, Haykin (1999) اختيار المعلمة C يساوي مجموعة قيم المخرجات. ان هذا الاقتراح منطقي، ولكن لا تضع في الاعتبار التأثير المحتمل للمتقدمين في بيانات التدريب.
- اقترح العديد من الباحثين منهم ((Hastie, Tibshirani, Friedman (2001) استخدام حساب احصائي خاص بنموذج (SVR)، إذ تقترن المعلمة (ϵ) باختيار دالة الخسارة، ويتم ضبطها على النحو الأمثل بكثافة ضوضاء معينة، في حين تفسر المعلمة C على انها معلمة تنظيم تقليدية يتم حسابها باستخدام طرائق مختلفة مثل التحقق المتبادل (cross – validation).
- في عام (2001) اقترح الباحث (Kwok) قيم مثالية لـ ϵ مقارنة الى حد كبير بالنسبة لتباين الضوضاء، بالاتفاق مع مصادر عامة بشأن المتجه الداعم، الا ان هذه الأساليب لا تعكس حجم العينة. [46]

4.3.2: تحليل المعلمات في انحدار المتجه الداعم (SVR)

من المهم في تصميم انحدار المتجه الداعم (SVR) اختيار المعلمة المثلى، وتكون على وفق الاتي:

1. معلمة التنظيم (C): تحدد C تكلفة المقايضة بين خطأ التدريب وتقليل تعقيد النموذج، كلما كانت درجة C أكبر، زادت كفاءه انحدار المتجه الداعم (SVR)، ولكن اذا كانت C كبيرة جدا (ما لا نهاية) فإن الهدف هو تقليل المخاطر التجريبية فقط دون اعتبار لتعقيد النموذج في جزء التحسين. [58]

$$\text{Min } R(y) = C \frac{1}{N} \sum_{i=1}^n L_{\varepsilon}(d_i, y_i) + \frac{1}{2} \|\omega\|^2 \dots (2.59)$$

2. معلمة تفاوت الخطأ (ε): تتحكم هذه المعلمة في عرض المنطقة (ε - insensitive) التي يتم استخدامها لملاءمة بيانات التدريب، فقد يتأثر عدد متجهات الدعم المستخدمة في بناء دالة الانحدار بقيمة (ε)، إذ ان القيمة الكبيرة لـ ε تؤدي الى اختيار عدد اقل من متجهات الدعم، مما تنتج عنه تقديرات انحداريه مسطحة، كما يؤثر في خطأ التقريب وخطأ التعميم.

3. معلمة σ: لبناء دالة (Kernel) يتم استخدام معلمة النواة (σ)، وتسهم معلمات (Kernel)

المختلفة في بنية الدوال، تستخدم الدالة (Gaussian) كدالة نواة لانحدار المتجه الداعم

(SVR)، وتشير معلمة (Kernel) الى عرض نطاق RBF .

تعتمد كفاءه وأداء انحدار المتجه الداعم (SVR) على التكوين الصحيح للمعلمات الثلاث

(C, ε, δ)، ففي كثير من الأحيان عند وصول كل معلمة الى نقطة مثلى لا يؤدي الى نتائج جيدة

لانحدار المتجه الداعم (SVR)، بحيث لا يمكننا الحصول على أداء جيد الا عندما يصل الجمع

بينها الى قيمة مثلى. [124] فيعتمد استقرار ودقة نموذج (SVR) الى حد كبير على النواة

الأساسية ومعلماتها (C, ε, σ)، فلا يتم بناء نموذج قوي وعالي الفعالية الا بعد اختياره هذه

المعلمات بعناية. [42] إذ ان اختيار المعلمات الثلاث (C, σ, ε) لنموذج انحدار المتجه الداعم

يؤثر في دقة التنبؤ، حيث ان (ε) تمثل معلمة دالة الخسارة والتي تتحكم في عرض دالة الخسارة

غير الحساسية (ϵ) (مثل انخفاض الضجيج في مجموعة البيانات) ومن ثم فإنه يحدد مقدار المتجه الداعم. بمعنى إذا كان ϵ كبير جداً فإن مقدار المتجه الداعم سيكون قليلاً وهذا يؤدي إلى أن نموذج التنبؤ بسيط وذو دقة منخفضة، أما إذا كان ϵ صغيراً جداً فيؤدي إلى تحسين دقة الانحدار، مما سيكون نموذج التنبؤ معقداً نسبياً. وتمثل المعلمة (C) معلمة التنظيم، وهي تحدد مدى تعقيد واستقرار نموذج التنبؤ. إذا كانت C كبيرة جداً فهذا سيحسن من دقة التعلم وعلى العكس من ذلك إذا كانت C صغيرة جداً فإن أخطاء التدريب تكون كبيرة.

يمكن الحصول على نموذج تنبؤ أكثر دقة وأكثر استقراراً عند تحديد المعلمة المناسبة لـ (C, ϵ) أما المعلمة (σ) فهي لا توضح القدرة الأساسية لدالة النواة فقط وإنما تعكس الارتباطات بين متجهات الداعم فمثلاً إذا كانت σ كبيرة جداً، يدل هذا على أن الارتباط بين متجهات الدعم قوية جداً ولا يمكن الحصول على دقة كافية. أما إذا كانت σ صغيرة فيدل هذا على أن الارتباط بين متجهات الدعم ضعيفة ومن ثم تكون عملية التعلم الآلي معقدة نسبياً، لذلك في عمليات النمذجة إذا كانت σ تقترب من الأصغر فيقترح تعيين قيمة أكبر لـ C . لذلك يتم في دراستنا الحالية اقتراح خوارزمية جديدة لتحسين تركيبة المعلمة، ولتحسين دقة التنبؤ لنموذج انحدار المتجه الداعم (SVR) والتي يتم تناولها لاحقاً. [88]

5.3.2: خصائص انحدار المتجه الداعم (SVR)

يتميز انحدار المتجه الداعم بعده خصائص يمكن إدراجها على وفق الآتي:

1. تحويل مقياس المشكلة: يتم إيجاد دالة القرار $f(x)$ من مجموع التدريب، حيث أن x يمثل متجه البعد، أن الطريقة المباشرة هي حل المشكلة الأولية باستخدام المقياس المعتمد على n وكلما زادت n تزداد تكلفة الحساب بسرعه وهو ما يمثل (curse of dimensionality)، أن المشكلة التي تم حلها في المتجه الداعم تمثل المشكلة المزدوجة إذ نلاحظ أن عدد متغيرات

هذه المشكلة هو l (عدد نقاط التدريب). وبعبارة أخرى، ان حجم المشكلة المطلوب حلها لا علاقة له مع ابعاد مساحة الادخال.

2. التناثر: تدل دالة القرار على ان نقاط التدريب المقابلة للمكونات غير صفرية α_i^* للحل α_i للمشكلة المزدوجة تكون منطقية، بعبارة أخرى فان نقاط التدريب المقابلة للمتجه الداعم تسهم فقط في دالة القرار، اما نقاط التدريب المتبقية المتوافقة مع المتجه غير الداعم فأنها لا تسهم في أي شيء وهذا يعني انه عندما تكون مجموعة التدريب كبيرة جدا فان نسبة متجهات الدعم تكون صغيرة، ومعظم المعلمات تكون مساوية للصفر α_i^* وهذا يعكس حقيقة تباين تصنيف المتجه الداعم وهو امر مهم لحساب المشكلات واسعه النطاق.

3. توظيف النواه: ان تطبيق النواه Kernell بدلا من الخريطة ϕ يحقق الانتقال من التصنيف الخطي الى التصنيف غير الخطي لذا يجب اختيار النواة (\emptyset) بدلا من خريطة ϕ مما يسهل من عملية الحساب وذلك لان حساب المنتجات الداخلية في المساحات عالية الابعاد يكون ذا تكلفة كبيرة. [53]

الفصل الثالث

نظرية الفوضى والخوارزميات الفوضوية الهجينة

1.3: مفهوم الفوضى

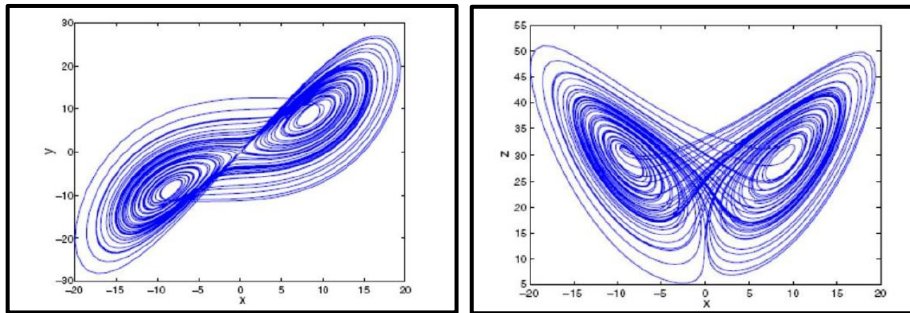
2.3: الخوارزميات الفوضوية الهجينة

3.3: النمذجة والتقلبات والتنبؤ

الفصل الثالث: نظرية الفوضى (Chaos theory) والخوارزميات الفوضوية الهجينة (Chaotic Hybrid algorithm)

1.3: مفهوم الفوضى

بدأت دراسة الفوضى من قبل خبير الأرصاد الجوي (Edward Lorenz) عام 1960 اذ استخدمها في التنبؤ بالطقس، وكان (Lorenz) ذا خلفية رياضية اذ يعتقد باستخدام الرياضيات في التنبؤ بتغيرات الطقس في محاولة لتقليد سلوكيات الطقس. وقد اكتشف (Lorenz) نموذجا حاسوبيا يستخدم اثنتي عشره معادلة محددة ومختلفة للحصول على النتائج نفسها وبالشروط المسبقة. نجح (Lorenz) في انشاء نموذج يناسب طريقة ملاحظته للطقس، فظهرت اختلافات فضلا عن وجود اضطراب منظم والهدف من وراء ذلك استخدام الحاسوب في محاكاة ظروف الطقس المتغيرة لاستقراء الظروف الجوية بشكل جيد، واستخدم معادلات مختلفة ليحاكي الاحوال الجوية وتأثيراتها عند تفاعل هذه الظروف فيما بينها، كما أراد (Lorenz) مراجعة بعض المخططات البيانية بشكل تفصيلي أكثر، فكرر دورة الحاسوب، ولاحظ اختلاف النتائج كل مرة عن سابقتها، على الرغم من ان البيانات السابقة هي نفسها، واستنتج ان تغييرا صغيرا في المعطيات الأولى لأحوال الجو ينتج عنه تغييرات ضخمة في أحوال الطقس الجوية، وهذا ما تم تأكيده لاحقا من قبل خبراء الفوضى فيشير مصطلح الفوضى الى العشوائية، وفروقات صغيرة جدا في الظروف الأولى للنظم يمكن ان تتضخم وتؤدي الى فروقات كبيرة جدا مع مرور الزمن. [15]



شكل (14) أسس نظام (Edward Lorenz) [98]

يمكن تعريف مصطلح الفوضى بمفهومها العام على انها تشويش او ضجيج او ارباك، في حين ان (Cornish) في عام 2007 عرف كلمة الفوضى بانها العشوائية وسلوك غير متوقع ولا يمكن التنبؤ به. [37]

1.1.3: نظرية الفوضى (Chaos theory)

تعد نظرية الفوضى من النظريات المثيرة جدا اذ انتشرت حديثا لتفسير الكثير من الأمور التي توقفت النظريات العلمية السابقة عن تفسيرها ولم تجد ما يبرر حدوثها ولا يزال هناك الكثير من الغموض حولها، فلاقت نظرية الفوضى اهتماما واسعا وأصبح بالإمكان إيجاد الصلة بين الأمور المتباينة او توقع الترابط بينها. تتبع الحركات مسارات غير خطية تتكرر وتتداخل بنسق معين غير متماثل تماما الا انها تكون منظمه جدا وكأنها تعود إلى نقطة جذب محددة بعد أن تتطلق عنها.

استمدت نظرية الفوضى جذورها من دراسة النظم اللاخطية لكونها محددة الخواص، اذ تهتم بالأنظمة العشوائية المعقدة وتحويلها الى أنظمة معلومة وملموسة من خلال مزج علم الرياضيات والهندسة الميكانيكية بعلم الاقتصاد، وارتكز اهتمام العلماء والباحثين على قدرة هذه النظم في تفسير الأشياء الغريبة التي تكون عشوائية وتشكل ضوضاء ولا يمكن توضيحها بالانماذج التقريبية. [15]

فنظرية الفوضى ظاهره غير خطية وغير منتظمة في العالم الطبيعي وهي حركة متقلبة للغاية وغير مستقرة في مساحة محدودة من الأنظمة القطعية، اذ يظهر النظام اعتماد استجابته على الظروف الأولية وله عدد لا حصر له من الردود الدورية المختلفة، عندئذ يقال للنظام غير الخطي فوضوي.

[118] وعرف (Stephen Kellert) نظرية الفوضى بانها دراسة نوعية السلوك غير المستقر

واللادوري في النظم الديناميكية الحتمية واللاخطية، ويعرف كل من (Fitzgerald and Eijnatten)

نظرية الفوضى بانها علم يصف ويشرح سلوك الأنظمة المعقدة او الديناميكية او غير الخطية بعيدا

عن أنظمة التوازن. [37]

وعلى وفق التعريفات المذكورة آنفاً لنظرية الفوضى، يمكن تحديد بعض الجوانب الأساسية للنظرية وهي:

1. سلوك النظام يكون غير مستقر ولا دوري أي بمعنى انه لا يكرر نفسه.
2. النظام مقيد وليس عشوائي، وتوقعاته لها شروط معينة مع مرور الوقت فسيتم تغيير إيجاد الديناميكيات مع مرور الزمن، بمعنى انه لا يمكن التنبؤ به على المدى الطويل والتوقعات تكون قصيره جدا.
3. كون النظام لا خطياً، فانه يعتمد على الحساسية للظروف الأولية أي ان مخرجات النظام لا تتناسب مع مدخلاته لذا ستؤدي التغييرات الطفيفة في الظروف الأولية الى نتائج مختلفة، كما تؤثر بشكل كبير في النظام قيد الدراسة. [37]

كثيرا ما تلاحظ ظاهره الفوضى في الأنظمة المعقدة والتي تمثل بسهولة من خلال تعريف تسلسل الفوضى اللوجستي وعلى وفق الآتي: [112]

$$x_i^{(n+1)} = \mu x_i^{(n)} (1 - x_i^{(n)}) \dots (3.1)$$

$$0 \leq x^{(i)} \leq 1 \quad , \quad i = (1,2,3, \dots, m)$$

حيث ان:

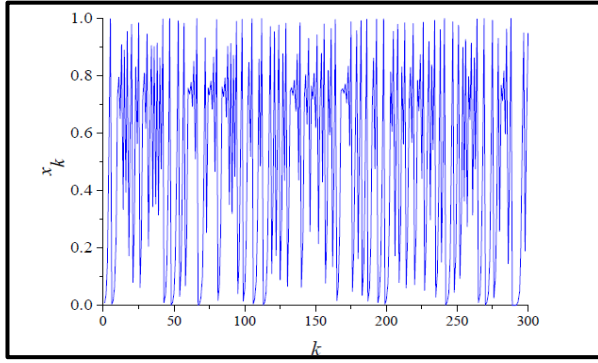
$$x^{(i)}: \text{تمثل قيمة المتغير الفوضوي } x \quad ; \quad 0 \leq \mu \leq 4$$

$$\mu: \text{تمثل معلمة التحكم} \quad ; \quad 0 \leq x^{(i)} \leq 1 \quad ; \quad \text{يمثل عدد تسلسل المتغيرات الفوضوية.}$$

m: يمثل العدد الإجمالي للمتغيرات الفوضوية.

يتم تحديد قيمه μ مع x ويستقر عند حجم ثابت وتتأرجح بين تسلسل محدود من الأحجام، ويعمل x بشكل فوضوي في نمط لا يمكن التنبؤ به. عندما $\mu = 4$ فان النظام يكون في حالة الفوضى، ومن ثمَّ فان السلوك الفوضوي يكون على النحو التالي:

$$x^{(i+1)} = 4x^{(i)}(1 - x^{(i)}) \dots (3.2) \quad ; \quad i = 1,2,3, \dots, k$$



الشكل (15) ديناميكيات الخريطة اللوجستية [118]

يوضح الشكل (17) الخصائص الديناميكية الفوضوية حيث ان $x_0 = 0.01$ ، والقيمة القصوى

$m=300$ ، نلاحظ ان هناك فرقاً صغيراً جداً في القيمة الابتدائية لمتغير الفوضى (x) والذي يسبب

فرقا كبيرا في سلوكها على المدى الطويل وتمثل السمة الأساسية للفوضى. [118]

يمكن تحويل المتغيرات الفوضوية إلى متغيرات التحسين من خلال استخدام التعيين الخطي وعلى وفق الآتي:

$$y_n = a + (b - a).x_n \dots (3.3)$$

حيث ان:

a : يمثل الحد الأدنى لمتغيرات التحسين ; b : يمثل الحد الأعلى لمتغيرات التحسين. [132]

فالمفهوم الرئيس لنظرية الفوضى هو أن تعمل الأنظمة بطرائق غير خطية وغير متوقعة بسبب

السلوك العشوائي، والمفهوم الأكثر شهرة لنظرية الفوضى حساسية النظام في الحالات الأولية، اذ

يشار إليه أيضا باسم تأثير الفراشة، ويعني هذا الافتراض أن شيئا بسيطا وغير ذي أهمية يكون له

تأثير أكبر بكثير ويمكن أن يؤدي إلى سلسلة من الأحداث المؤدية لحدث واسع النطاق، وهذا

الافتراض أساس الظاهرة المعروفة (بأثر الفراشة). [59]

لفهم المتغيرات تستخدم النظرية لإيجاد حل للمشكلات غير الخطية الا ان الاستنتاجات لا يمكن

التنبؤ بها، اذ ان هناك العديد من المشكلات في العالم الحقيقي، مثل عدم استقرار الطقس، ونشاط

سوق الأوراق المالية وما إلى ذلك، هي مشكلات غير خطية تكون حالتها غير مستقرة في وقت

محدد ففي الحل النهائي لمشكلة ما تنص النظرية على إن حدوث خلل طفيف للتغيرات في مجال

الإدخال سيؤدي إلى تغييرات جذرية ومتغيرة. تستخدم هذه النظرية لتحديد الحلول العالمية المثلى للخوارزميات الوراثية على أساس هذا المفهوم من عدم القدرة على التنبؤ وخط المتغيرات. [109]

2.1.3: عناصر نظرية الفوضى

تتكون نظرية الفوضى من ثلاثة مكونات رئيسة وهي كالاتي:

1.2.1.3 - التنظيم الذاتي (Self Organization) [37]

هي عملية تنظيم مجموعة من القواعد البسيطة في شبكة مترابطة، لها القدرة في الحصول على المعلومات التي تولدها بنفسها واستخدامها كمدخلات وهو ما يمثل قدره النظام الفوضوي على تنظيم نفسه. يشير مصطلح التنظيم الذاتي الى مفهوم الأنظمة التكيفية المعقدة لأنها تتمكن من التطور او التحول او التكيف مع البيئة أو التفاعل معها. يرجع السلوك الديناميكي الذي يمتاز به نظم التنظيم الذاتي الى السلوكيات التي تتفاعل في أجزاء النظام.

كما أشار Frederick في عام 1998 الى ان التنظيم الذاتي يشير الى قدره النظام على التكيف مع الظروف المحيطة به من اجل البقاء على قيد الحياه، تحدث عملية إعادة التنظيم الذاتي نتيجة التغيرات في البيئة الخارجية او الاختلاف الجزئي الذي يؤول الى توسيع النظام. فإذا كانت الاستجابة للتغيرات البيئية جيدة فان إعادة التنظيم الذاتي الجديدة ستكون متوافقة مع هذه التغيرات. فالفشل في هذه الأنظمة هو عدم القدرة على التعامل بشكل أفضل مع متغيرات البيئة.

2.2.1.3 - التشعبات (Bifurcation)

تعني التشعبات تغير نوعي في سلوك نظام ديناميكي معقد وفوضوي غير مستقر في بيئته بسبب الاضطراب او الضغط. فستوجه الحالة الجاذبة مسار الضغط الى ما يسمى انتقال الطور في نقطة زمنية معينة، اذ يتشعب النظام ويدخل في حالة النظام الجديد من خلال التنظيم الذاتي او الدخول في حالة من الفوضى. يدخل النظام مرحلة التشعب في مجال افتراضي، قد يختار النظام أكثر عوامل الجذب نفوذاً، او قد يقفز من جاذب الى آخر. في هذه المرحلة يتبنى النظام الخيارات

المستقبلية، يطلق على هذه المرحلة بالفوضى العميقة، او يعيد النظام ترتيب نفسه بدرجة اعلى من التعقيد، أو تتلاشى وتختفي وتسمى هذه العملية حافة الفوضى (edge of chaos). وللتشعبات ثلاثة أنواع وهي كالآتي:

- كارثي (Catastrophic): في هذه الحالة يكون الانتقال حادا مما يسبب المزيد من الاضطراب.

- متفجر (exploding): يكون الانتقال في هذه الحالة مفاجئا وتتحكم فيه عوامل متقطعة تجعل المنظومة تنتقل من نظام الى اخر.

- خفي (subtle): يكون الانتقال فيه سلسا. [96]

3.2.1.3 - الجاذب (attractor)

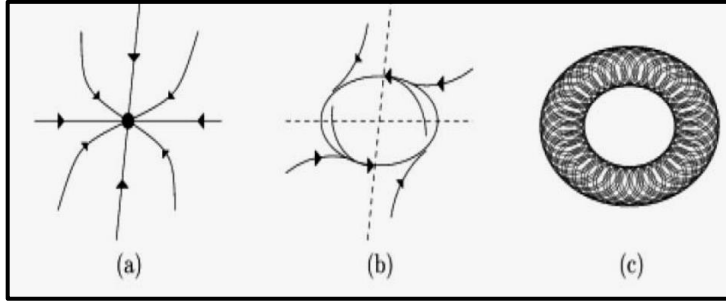
عندما يكون هناك نظام غير خطي في مرحلة الفضاء الفوضوي يتم تشكيل نمط سلوكي وهو ما يعرف بالجاذب، أي يكون مرتبطاً بنظام السلوك العشوائي المضطرب. مما يمكن النظام الفوضوي في مجالين هما (الاستقرار وعدم الاستقرار) وضمن نطاق معين.

كما تم تعريف الجاذب على انه مجموعة محدودة من المسارات التي تمثل المسار الذي يتبعه سلوك المجتمع. يميل سلوك النظام غير الخطي الى الانكماش والتقلص في مناطق محددة ويدعى هذا الانكماش بالجاذبية، لأنه يمثل في الواقع مجموعة من النقاط تتلاقى معها جميع المسارات.

تحتوي المنظومات على ثلاثة أنواع من الجاذب وهي كالآتي:

- الجاذب الدوري: يقوم هذا النوع بتكرار ذاته في تأرجحات دورية.
- الجاذب النقطي: يستخدم هذا النوع في المنظومات التي تصل الى توازن مستقر، ثم يعود ليتوقف في نقطة واحدة.

- الجاذب الغريب: يوجد هذا النوع من الجاذب في المنظومات التي تصل الى مرحلة حافة الفوضى بشكل غير قابل للتنبؤ به [37]، ويبين الشكل (18) حركة الجاذب الثلاثة:



(a) شبه دوري (b) الدوري (c) نقطي

الشكل (16) يمثل الأنواع الثلاثة من الجواذب. [37]

3.1.3: الخصائص الأساسية للأنظمة الفوضوية

تحدث الفوضى في بعض النظم القطعية غير الخطية، اذ تظهر الفوضى عندما يكون هناك تطور

طويل الأجل المستمر وغير المنضبط والذي يفى بمعايير رياضية معينة، ف لوحظت في أنظمة

الفوضى مجموعة من الخصائص تم تلخيص أهمها على وفق الاتي:

1. عدم الاستقرار الديناميكي (Dynamic instability): تعد خاصية الحساسية للظروف

الأولية من أشهر مميزات النظم الفوضوية. وغالبا ما يشار اليه أيضا باسم تأثير الفراشة اذ

يتطور شرطان اوليان مغلقان بمسارات مختلفة ومتباعدة بشكل كبير.

2. الخلط الطوبولوجي (Topological mixing): بديهيا يتم تصويره على انه خلط الاصباغ

الملونة، يعني ان النظام سيتطور في الوقت المناسب وغالبا ما يتم تحويله او تتداخل أي

منطقة معينة من الدول مع أي منطقته معينة أخرى. [44]

3. اللا دورية (Aperiodicity): تعد الميزة الرئيسية للأنظمة الفوضوية، اذ تعرف هذه الخاصية

بانها خاصية نظام ذي تذبذبات غير منتظمة أي لا زيادة ولا نقصان ولا ثبات، بعبارة أخرى ان

هذه الآلية ليست عشوائية ولا دورية. [99]

4. مدارات دورية كثيفة (Dense periodic orbits): يتبع النظام ديناميكية يمكنها الاقتراب

من كل حالة مقارنة ممكنة.

5. تشوهه (Ergodicity): تعطي القياسات الإحصائية للمتغيرات نتائج مماثلة بغض النظر عما إذا

تم اجراؤها عبر الزمن أو المكان. بعبارة أخرى، تظهر الديناميكيات إحصاءات مماثلة عند قياسها عبر الزمن أو المكان.

6. التشابه الذاتي (Self – similarity): تطور النظام في الزمان أو المكان، أي يظهر

المظهر نفسه على مستويات مختلفة من الملاحظة، هذه الخاصية تجعل النظام يظهر بشكل

تلقائي ومتكرر في مستويات مختلفة من الملاحظة. [44]

2.3: الخوارزميات الفوضوية الهجينة (Hybrid chaotic Algorithms)

اكتشف العديد من الباحثين مؤخرًا استخدام التهجين في مسائل الأمثلية يمكن أن يحسن نوعية

الحلول التي يمكن إيجادها، على الرغم من وجود اختلافات كبيرة ومهمة بين هذه الخوارزميات، إلا

أنها تشترك في بعض العناصر التي تم تجاهلها وجعلها غير مستخدمة. [9]

يتم تطبيق الخوارزميات لحل مشكلات التحسين وتحديد القيم المناسبة لمعاملات نموذج انحدار

المتجه الداعم (SVR) لاستخدامها في عملية التنبؤ، وسيتم توضيح ثلاثة أنواع من الخوارزميات

لكثرة استخدامها من قبل الإحصائيين وملاءمتها مع الأساليب الإحصائية.

1.2.3: الخوارزمية الجينية الفوضوية (chaotic genetic Algorithm) واستخدامها في

تحديد معاملات نموذج انحدار المتجه الداعم (SVR)

قبل التطرق إلى الخوارزمية الجينية الفوضوية لآبد من توضيح الخوارزمية الجينية في هذه الفقرة

1.1.2.3: الخوارزمية الجينية (genetic algorithm)

تعد الخوارزمية الجينية (GA) إحدى طرائق الذكاء الاصطناعي، لأهميتها في حل المشكلات

المعقدة، إذ أنها تحتوي على عدد كبير من الحلول البديلة، كما أن الحل الناتج من تطبيق

الخوارزمية الجينية (GA) يكون في معظم الحالات قريب من الحل الأمثل. تعتمد الخوارزمية

الجينية على الانتقاء الطبيعي وأنظمة الجينات الطبيعية. وتستند فكرة الخوارزمية الجينية (GA)

إلى فكره الهندسة الوراثية التي تتميز بالإنتاج المقصود لمجموعات موروثية لتكوين أفراد يتمتعون

بصفات جيدة، وعلى هذا الأساس تختار الخوارزمية الجينية الحل الأفضل من بين عدد كبير من الحلول واجراء بعض التداخلات والتغييرات بين هذه الحلول لإنشاء أفضل حل. [97]

تم تطوير استخدام هذه الخوارزمية لأول مرة من قبل العالم جون هولاند في عام 1975 في جامعة Michigan لتمثيل فئة فرعية من الخوارزميات التطورية، تقوم الخوارزميات الجينية بحل مشكلات التحسين من خلال استخدام فكرة التطور الطبيعي، فهي تولد أفراداً جديداً مع عوامل الاختيار، العبور والطفرات. [38]

تم تطبيق الخوارزمية الجينية بنجاح لإيجاد حلول مقبولة (قريبة من المثالية) في المشكلات المتعلقة بالعلوم (بما في ذلك العلوم الطبية والهندسية) لأنها تقلل الكثير من الوقت والجهد لمصممي الأنظمة والبرامج عن طريق إيجاد خوارزمية عامة يعتمد عليها عند حل أنواع مختلفة من المشكلات، بدلا من بناء خوارزمية خاصة لكل مشكلة، والنظر في الحجم ونوع البيانات المستخدمة وطبيعة دالة الهدف وقيود كل مشكلة مع مراعاة التغييرات اللازمة التي تتناسب مع خصوصية كل مشكلة. [6]

تستخدم الخوارزمية الجينية (GA) لتحسين المعلمات في نمذجة انحدار المتجه الداعم (SVR) على نطاق واسع لإيجاد الحل الأمثل، تعمل الخوارزمية الجينية مع المكونات الأساسية الآتية:

1. الفرد (The individual): عند التعامل مع الخوارزمية الجينية تكمن الصعوبة الرئيسية في

إيجاد طريقة لتمثيل البيانات المعروفة باسم مخطط الترميز. لذا يجب تحويل المشكلة الى

سلاسل فردية لحلها باستخدام الخوارزمية الجينية (GA).

2. اللياقة البدنية للتحسين (Fitness function for optimization): تعد اهم أجزاء الخوارزمية التي تحاول تحسينها.

3. دالة الهدف (The objective function): يحصل كل فرد على معدل او معيار يسمى

اللياقة بعد عملية التقييم، قد يتمتع بعض الافراد باللياقة البدنية الجيدة لذا فانهم اقرب الى الحل

المطلوب الوصول اليه، اولديهم لياقة بدنية رديئة أي يبتعدون عن الحل. وتتوقف الخوارزمية عندما يكون شخص او أكثر قد حقق درجة مناسبة من اللياقة أي عندما يتم العثور على الحلول التي تم البحث عنها.

4. مجموعة من السلاسل (الكروموسومات) (Population of chromosomes): السلسلة

تشير الى القيم المتمثلة بالحل المرشح للمشكلة المراد حلها. تبدأ الخوارزمية الجينية بمجموعة حلول مختاره عشوائيا وتعمل كمجموعة أولية من السكان، ثم يتم تقييم كل حل من قبل دالة اللياقة البدنية لاختبار مدى نجاحها في حل المشكلة.

5. السكان (The population): يمثل السكان مجموعة من الافراد أي مجموعة من الحلول

الممكنة، اذ ان الجيل يعكس حالة السكان عند تكرار الخوارزمية.[61]

ان اهم مفاهيم المجتمع هي الجيل الابتدائي وحجم المجتمع (n). يرتبط حجم المجتمع بدرجة تعقيد المشكلة، لأن طبيعة المشكلة تختلف من نوع لآخر ويتم الاختيار حسب تقدير الباحث، في بعض الأحيان يمكن ان يتسبب حجم المجتمع في حدوث المشكلة، فكلما كان الحجم أكبر كان من الأسهل استخدام نطاق البحث بأكمله. ومع ذلك يتم توفير الوقت المطلوب على وفق العلاقة الاتية:

$$n \log n \dots (3.4)$$

تكون العلاقة طردية، اذ تزداد بزيادة الحجم.

اثبت العالم Goldberg في عام 1989 ان فعالية الخوارزمية الجينية تتأثر بحجم السكان، وتعود الأفضلية الى الحجم الأكبر لكنها تكون مكلفة من حيث السرعة والأداء.

يمثل الجيل الابتدائي مجموعة افراد المجتمع في بداية عملية البحث يجب ان يتضمن الجيل الابتدائي مجالا واسعا ومتنوعا لتغطية نطاق البحث بشكل أفضل، لذلك يتم اختيار الجيل الابتدائي بشكل عشوائي، ومن الممكن استخدام طرائق الاستدلال عند اختيار الافراد مما يؤدي الى أداء أسرع. [22]

6. التقاطع (Crossover): تستخدم لإنتاج الحل التالي من السلاسل الفردية.

7. طفره تلقائية من السلاسل الفردية: في الحل الجديد تقلب (bits) الفردية في الحلول الجديدة

بشكل عشوائي.

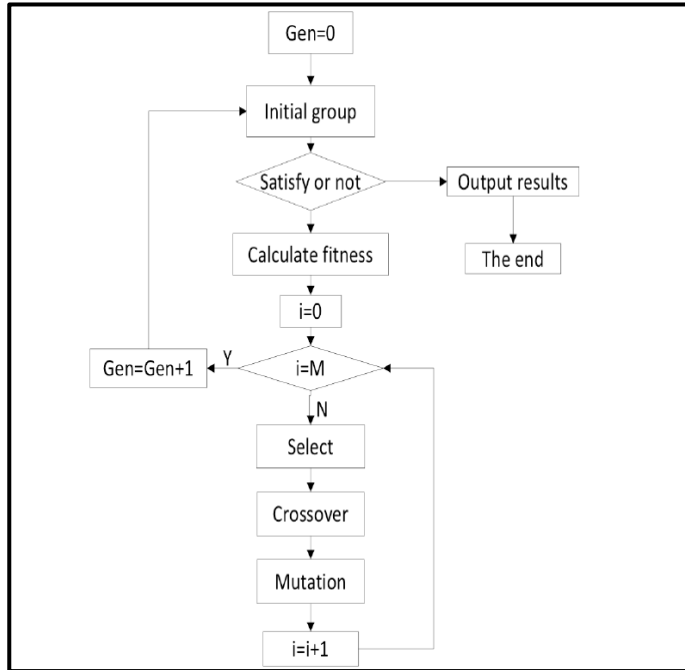
استخدمت الخوارزمية الجينية (GA) على نطاق واسع في مجال التنبؤ بالحمل الكهربائي، إذ تكون مناسبة تماما للأنظمة غير الخطية. تعتمد عملية التحسين على الاختيار الطبيعي لأفضل الحلول التي يتم التوصل إليها من مجموعة واسعة من سكان (نموذج التنبؤ). يتم نشر هذا النوع من التحسين المستند إلى الخوارزمية الجينية (GA) عادة في أثناء إجراء اختيار النموذج أي عند العثور على المعلمات الأكثر ملاءمة لنموذج التنبؤ [73].

يتم تطوير مجموعة من الحلول المرشحة بواسطة الخوارزمية الجينية (GA) ويمثل كل حل بواسطة سلسلة تسمى (كروموسوم)، ويتم عادة ترميزها كسلسلة ثنائية. يكون إجراء الخوارزمية الجينية (GA) بشكل تكراري بحيث يتم الاحتفاظ بحل مرشح ثم تقوم بفرز إيجابياته وسلبياته ثم اختيار بعض من الحل على وفق مقاييس معينة واستخدام العوامل الوراثية لتوليد جيل جديد من الحلول المرشحة. يتم تكرار هذا الإجراء الى ان يتم التوصل إلى تقارب الهدف. بعد ذلك، يتم ترميز الفرد بأرقام حقيقية وتمثيل كل فرد بواسطة كروموسوم وحساب اللياقة البدنية لكل فرد، وتتطور الأفراد نحو الحل الأمثل على مدى أجيال متعاقبة عن طريق الاختيار والعبور وعمليات الطفرات، وتستمر عملية التطور الى ان يتم الوصول إلى معايير الإيقاف.

1.1.1.2.3: طريقة عمل الخوارزمية الجينية (Genetic Algorithm)

تحتوي الخوارزمية الجينية (GA) على العديد من الخطوات الأساسية لصياغة الحل الأمثل لمشكلة معينة، وتكون ثابتة لمختلف المشكلات والتطبيقات، ويكمن الاختلاف في صياغة كل خطوة من الخطوات وتطبيقها حسب المشكلة أو مجال تطبيقها، تكون هذه الخطوات مترابطة بعضها مع بعض ولا يمكن تطبيق الخوارزمية على أي مشكلة ما لم يتم تطبيق جميع هذه الخطوات والا فان الخوارزمية الجينية ستفقد قيمتها وفائدتها لإيجاد الحلول أو تحسينها، [11] وتظهر إليه عمل

الخوارزمية الجينية على وفق مخطط (1)



مخطط (1) يمثل المخطط الانسيابي للخوارزمية الجينية (GA)[77]

1.1.1.1.2.3- التهيئة (Initialization)

في البداية يتم إنشاء عدة حلول فردية بشكل عشوائي لتشكيل عدد السكان الاولي. فيتم تحديد حجم السكان من خلال طبيعة المشكلة، ولكنه يتطلب عادة عدة مئات أو آلاف من الحلول المحتملة. تقليدياً، فيتم إنشاء السكان بشكل عشوائي ليشمل جميع النتائج الممكنة (مساحة البحث). في بعض الأحيان، من المرجح أن يتم اكتشاف الحلول المثلى، ويمكن أن تكون الحلول "مصنفة".

2.1.1.1.2.3- الانتقاء او الاختيار (selection)

يكون اختيار جزء من السكان الحاليين خلال كل جيل متتال لتكاثر جيل جديد، فيتم اختيار الحلول الفردية اعتماداً على اللياقة البدنية، إذ يتم اختيار الحلول المجربة (على النحو الذي تحدده دالة اللياقة البدنية)، وتقوم طرائق اختيار معدل اللياقة البدنية بتقييم مدى ملاءمة كل حل واختيار أفضل الحلول بشكل تفضيلي. وبما أن هذا النهج يمكن أن يستغرق وقتاً طويلاً، فإن هناك أساليب أخرى تقوم بتقييم عينة عشوائية فقط من السكان. أما الحلول ذات قيم اللياقة البدنية العالية فأنها تعطي حلاً أفضل في الجيل التالي استناداً الى دالة اللياقة البدنية، ولعملية الاختيار تطبق نظرية

اختيار عجلة الروليت. ان معظم الدوال تكون عشوائية ومنظمة لاختيار نسبة صغيرة من الحلول التي هي أقل ملاءمة. ويساعد هذا في الحفاظ على تنوع كبير في السكان مما يقلل من التقارب على الحلول الضعيفة، [40] ويتم توضيح استراتيجيات الاختيار على النحو الآتي:

1. اختيار العجلة المتدرجة (roulette wheel selection)

تعد أحد الطرق التقليدية في الخوارزمية الجينية، يركز مبدؤها على الاختيار على وفق العجلة، إذ يتم تقسيم العجلة بشكل عشوائي الى قطاعات لتقييمها باحتمالية اختيار الفرد على وفق دالة اللياقة البدنية. [22] ويعني احتمال اختيار الافراد بأنه نسبة قيمة اللياقة البدنية الى اجمالي قيم اللياقة البدنية لجميع الحلول (الكروموسومات)، أي أن احتمالات اختيار الأفراد تتناسب خطياً مع قيم لياقتهم باستخدام هذه الطريقة. [87]

يتم حساب هذه الاحتمالات على وفق الصيغة الآتية:

$$p(i) = \frac{f(i)}{\sum_{i=1}^n f(i)} \dots (3.5)$$

حيث ان:

$p(i)$: احتمال اختيار الفرد (i) ; $f(i)$: يمثل قيم اللياقة البدنية للفرد (i) ; n : يمثل عدد

الافراد. [5]

ويتم اجراء هذه العملية على وفق الآتي:

- استخدام معيار اللياقة البدنية لحساب الاحتمالات التراكمية للفرد $C(i)$ مع كل حل على حدة وعلى وفق الصيغة الآتية:

$$C(i) = \sum_{j=1}^i P(j) \dots (3.6)$$

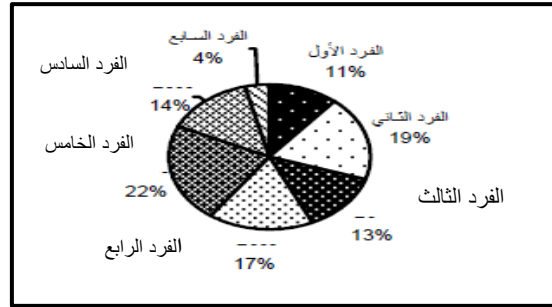
- يتم أولاً حساب اللياقة البدنية لجميع المدخلات وتمثيلها على العجلة بنسبة مئوية.
- يتم تدوير العجلة المتدرجة في مساحة البحث عن الكروموسومات "N".
- يتم اختيار حل (كروموسوم) ذي لياقة أكبر لعدة مرات. [22]، [40]

ويوضح الجدول (1) مثالا على اختيار العجلة المتدرجة

جدول (1) يمثل قيم العجلة المتدرجة. [40]

الكروموسوم	قيم اللياقة البدنية	احتمالية الاختيار	الاحتمال التراكمي
الفرد 1	13	0.11	11%
الفرد 2	22	0.19	30%
الفرد 3	15	0.13	43%
الفرد 4	19	0.17	60%
الفرد 5	25	0.22	82%
الفرد 6	16	0.14	96%
الفرد 7	5	0.04	100%

عند استخدام العجلة المتدرجة، فإن الحلول (الكروموسومات) التي تتمتع بأعلى درجات اللياقة لديها أعلى احتمالات الاختيار، في حين أن البقية لها أدنى المستويات. لذلك، غالبا ما يتم اختيار الافراد عالية الاحتمالية، وهذا ما يؤكد شكل (19)، وتتمثل اهم مزايا هذه الطريقة بتجنب التقارب السابق الى الحد الأقصى المحلي. [87]



الشكل (17) يمثل انتقاء عجلة الروليت. [5]

2. اختيار المسابقة (Tournament Selection)

هذا الاختبار يتضمن الآتي:

- مجموعة من الحلول المحتملة، يتم اختيار اثنين من الحلول ومقارنة لياقتهم البدنية مما يسمح بإعادة إنتاج الأفضل.
- يتم اختيار أفضل فرد في كل مسابقة وذلك من خلال اختيار المسابقة الحتمية.
- يمكن الاستفادة من تصميم متوازٍ.

3.1.1.1.2.3 - التقاطع (Crossover)

يعد العبور المشغل الأكثر أهمية في الخوارزمية الجينية إذ يختار هذا المشغل الجينات من الافراد

الأصل وينتج نسلا جديدا وأفضل طريقة للقيام بذلك هو تحديد نقطة التقاطع، إذ ينتج مشغل

العبور حلين جديدين بتبادل المعلومات عن الحلول المختارة. وتوجد طرائق أخرى لكيفية إجراء

التقاطع مثل اختيار المزيد من النقاط المتقاطعة. يتم تحسين أداء الخوارزمية الجينية عندما يكون

مشغل التقاطع معقداً جداً وبالإمكان اعتماده على ترميز الافراد. [40]

لا تنفذ عمليات العبور على كل فرد مختار، تختار الخوارزمية الجينية ما إذا كان سيتم تنفيذ عملية

العبور على زوج معين من الافراد، اعتمادا على احتمال معين يسمى باحتمال التقاطع الذي يتم

توليده من قبل المستخدمين. [87]

4.1.1.1.2.3 - الطفرة (Mutation)

تحدث الطفرات بعد إجراء عملية التقاطع لمنع وقوع جميع حلول السكان في المستوى المحلي عند

حل المشكلة، إذ تؤدي الطفرات دوراً ثانوياً بعد عملية العبور في الخوارزمية الجينية. فتغير عوامل

الطفرات بعض الأجزاء العشوائية المحددة من الحلول (الكروموسومات) فإذا كانت الحلول عبارة

عن سلاسل ثنائية سيتم استخدام الترميز ثنائي التبدل من 1 إلى 0 أو من 0 إلى 1 [41].

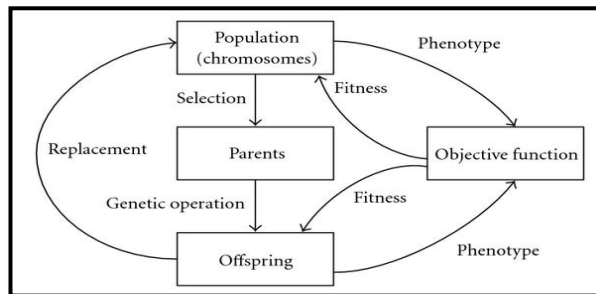
ويمكن أن تتبع الطفرات النحو التالي:

جدول (2) الطفرة. [40]

Original offspring 1	1101111000011110
Original offspring 2	1101100100110110
Mutated offspring 1	1100111000011110
Mutated offspring 2	1101101100110110

تعتمد الطفرة على كل من الترميز والتقاطع على سبيل المثال، قد تؤدي الطفرة الى تبادل جينين

عند استخدام ترميز التباديل.



الشكل (18) دورة الخوارزمية الجينية (GA). [40]

5.1.1.1.2.3 - مقياس التوقف (Stopping criteria)

تم اعتماد عدة مقاييس في الخوارزمية الجينية لتحقيق شرط التوقف في حالة عدم وجود حل مثالي.

بما في ذلك تنفيذ الخوارزمية الجينية وعدد الأجيال المطلوبة، وتحديد الوقت اللازم لتنفيذ الخوارزمية في المجتمع او القيمة المثلى لدالة اللياقة المطلوبة عندما تكون اقل من قيمة اللياقة التي تم تحديدها في حالة (Minimize)، وتتوقف عندما تكون القيمة المثلى لدالة اللياقة اكبر من قيمة اللياقة المحددة في (Maximize). يتم اختبار مقياس التوقف للخوارزمية الجينية بعد انشاء كل حل جديد لمعرفة ما إذا تم الوصول الى الحل المطلوب. [34]، [29]

2.1.1.2.3: دالة اللياقة البدنية (Fitness function) [34]

تعد واحدة من اهم المكونات الأساسية التي تنفذها الخوارزمية الجينية، فتحسب لياقة كل فرد، تسمى أحيانا بدالة الهدف للخوارزمية الجينية يتم من خلالها تحديد عدد الافراد الذين يتم اعتمادهم في حل المسائل. فاللياقة البدنية تعبر عن قدرة الفرد على التأقلم مع البيئة، ويتم استخدامها لتحديد لياقة كل فرد، يتطلب بناء دالة لياقة لكل فرد لتقييم الأداء قبل البحث عن القيم المثلى لمعاملات انحدار المتجه الداعم (SVR). من اجل تقييم واختيار الحلول (الكروموسومات) المناسبة، تتطلب عملية تحسين معاملات (SVR) من خلال الخوارزمية الجينية (GA) تحديد دالة اللياقة البدنية. وقد تم تطبيق مقاييس عديدة لهذا الغرض منها متوسط مربعات الخطأ (MSE)، فيستخدم لقياس جودة كل حل حتى يمكن معالجة المواد الوراثية الجيدة، كذلك يتم استخدام معيار متوسط نسبة الخطأ المطلق (MAPE) كدالة اللياقة البدنية. [42]

3.1.1.2.3: خصائص الخوارزمية الجينية (GA)

تتميز الخوارزمية الجينية (GA) بالخصائص التالية:

1. تكون العشوائية عامل رئيس في الخوارزمية الجينية (GA)، وذلك لان الانتقاء والطفرة والتكاثر تتطلب جميعها عملاً عشوائياً.
2. تبحث معظم الطرائق الكمية عن حل واحد في حين تبحث الخوارزمية الجينية عن مجموعة من الحلول، اذ تحتفظ بالحلول المتعددة وفي كل تكرار ستحصل على المزيد من الحلول. ويؤدي

هذا دورا مهما في متانة الخوارزمية الجينية لتوفر فرصة للشمول وتقادي التحيز.

3. عند استخدام الخوارزمية الجينية لا توجد شروط او متطلبات خاصة، ويمكن تطبيقها على أي مشكلة.

4. لا تستخدم الخوارزميات الجينية المشتقات، وانما تعتمد على دالة اللياقة (دالة الملاءمة)، فدالة

اللياقة يمكن ان تكون أي دالة قابله للتقييم أي لا يوجد نوع محدد من الدوال الرياضية، وتعد

هذه الخاصية الأكثر تميزا للخوارزمية الجينية.

5. تتعامل الخوارزمية الجينية مع المتغيرات الكمية المستمرة والمنفصلة، وعاده ما تستخدم مع

المتغيرات المرمزة.

كذلك فان الخوارزمية الجينية كغيرها من الطرائق تشوبها بعض العيوب والتي يتم ادراجها على وفق

الآتي: [22]

1. إذا فشلت جميع الوسائل والطرائق الشائعة في إيجاد حل او تقديم دعم إضافي، فان

الخوارزميات الجينية تستحق الاستخدام.

2. مبدأ عمل الخوارزمية الجينية عام جدا، يختلف من حيث الأداء والسرعة، ويتأثر بنوع المشكلة.

3. في بعض الحالات، قد تتم مواجهة بعض الغموض والصعوبات عند تحديد دالة الملاءمة بدقة،

او اختيار المعلمات الخاصة بعمليات الخوارزمية الجينية.

4. تقترح الخوارزمية الجينية حلا تقريبا جيدا ومقبولا بسبب التقارب السابق لأوانه ووقوعها في

المستوى الأمثل المحلي (local optimum)، أي لا تضمن أفضل حل للمشكلة.

ان الخوارزمية الجينية (GA) تكون مناسبة لحل المشكلات المعقدة وغير خطية والتي يصعب

تناولها في أساليب البحث التقليدية، ولها قوة أكبر من خوارزميات البحث التقليدية في البحث

العالمي، الا انه عندما تبحث الخوارزمية الجينية عن أفضل حل فان إمكاناتها ترمي الى

الانخفاض، واحتمال وقوعها في الحد الأدنى المحلي. لذا يتم دمج تسلسل الفوضى في الخوارزمية

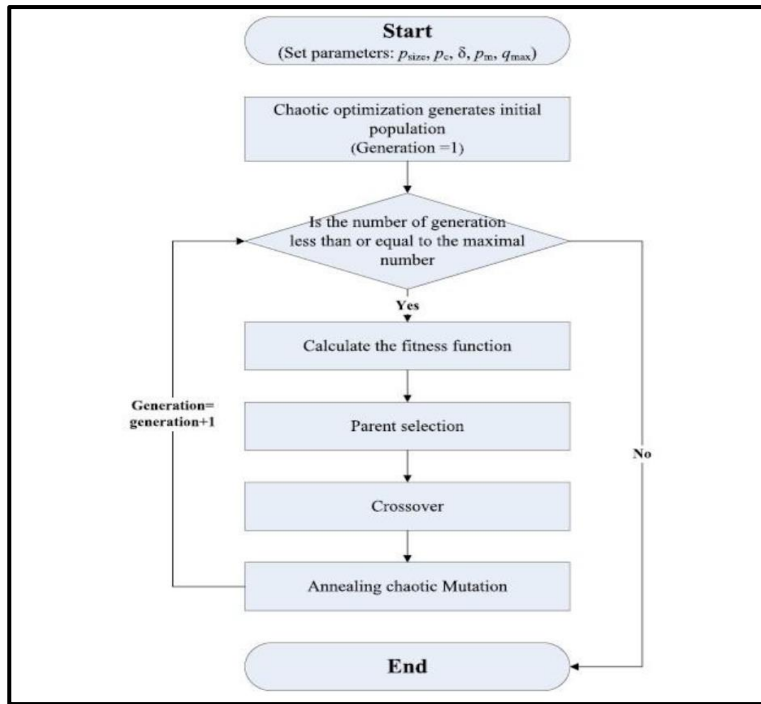
الجينية (GA) لحل هذه المشكلة، فالمبدأ الأساسي لهذه الخوارزمية هو تحميل المتغيرات الفوضوية والبحث عن الحل الأمثل من خلال عملية البحث الفوضوية، وتدعى هذه الخوارزمية الجديدة بالخوارزمية الجينية الفوضوية (CGA). [119]

2.1.2.3: الخوارزمية الجينية الفوضوية (Chaotic genetic algorithm)

نتيجة لوجود عيبين رئيسيين في الخوارزمية الجينية (GA) وهما بطء التقارب ووقوعها في المستوى الأمثل المحلي (local optimum)، إذ لا يمكن الاحتفاظ بعدد السكان الأصليين تحت ضغط انتقائي مما يؤدي الى انخفاض التنوع السكاني، وهكذا فإن العديد من الحلول الأولية ستكون ضعيفة وبعيدة عن المستوى الأمثل دون عمليات الاختيارات السليمة لاحتمالات الانتقال والطفرة. [69] كما انها تستغرق وقتا طويلا في التدريب لذا تم الجمع بين تسلسل الفوضى الهجين والخوارزمية الجينية (GA) للتغلب على هذه المشكلات.

فالخوارزمية الجينية الفوضوية (CGA) اداة قوية لحل مشكلات التحسين غير الخطية والمعقدة، والمبدأ الأساسي لها تحويل متغير المشكلة من فضاء الحل الى فضاء الفوضى ثم اجراء بحث لمعرفة الحل على وفق ثلاث خصائص، العشوائية (random city)، والتشوه (Ergodicity)، والانتظام (regularity) للمتغيرات الفوضوية. [70]

والمخطط (2) يوضح الشكل الانسيابي للخوارزمية الجينية الفوضوية



مخطط (2) الشكل الانسيابي للخوارزمية الجينية الفوضوية. [123]

1.2.1.2.3: ميزات الخوارزمية الجينية الفوضوية (Chaotic Genetic Algorithm)

ان طريقة عمل الخوارزمية الجينية الفوضوية (CGA) تشبه طريقة عمل الخوارزمية الجينية (GA)

ولكنها تختلف في طول سلاسل الحلول (String length)، اذ تكون غير ثابتة الطول.

تستخدم هذه الخوارزمية لحل المشكلات الصعبة، وهي أكثر كفاءه وفعالية وتعمل بنظام مرن

لتمثيل وحل مشكلات سلسلة الحلول، وتتكون السلسلة من جينات مختلفة من سلسلة لأخرى (سلسلة

الحلول)، إذ يقوم البرنامج بإنشاء سلسلة قصيرة تمثل حلاً جزئياً للمشكلة قيد الدراسة وتقييمها

(حساب لياقتها) في المراحل الأولى من الخوارزمية الجينية (GA) وتحتفظ بالجينات التي تحقق

معدل لياقة اعلى من المتوسط. ثم تبدأ الخوارزمية عملها الطبيعي بالانتقال من جيل الى اخر في

الفئة العمرية الأولى، وفي كل جيل تتم زيادة طول السلسلة الناتجة للوصول الى الطول الكامل

لسلاسل الحلول. [21]

2.2.1.2.3: استخدام الخوارزمية الجينية الفوضوية (CGA) في تحديد معاملات نموذج انحدار المتجه الداعم (SVR)

للتغلب على العيوب التي تعاني منها الخوارزمية الجينية (GA) تم اكتشاف بعض النهج والتحسينات الفعالة التي تحافظ على التنوع السكاني وتتجنب الوقوع في المستوى الأمثل المحلي، ومن هذه الحلول المحتملة تقسيم الكروموسومات الى عدة مجموعات فرعية والحد من الانتقال بين الأعضاء الى مجموعات فرعية منفصلة، إذ تتطلب هذه الطريقة عددا كبيرا من السكان. الا انه توجد حلول اخرى تركز على نهج الفوضى لسهولة تنفيذه وقدرته الخاصة على منع الوقوع في المستوى الأمثل المحلي. فتم اقتراح خوارزمية التحسين الفوضوي لحل وتحسين الدوال المعقدة استنادا الى نهج ومميزات الفوضى. [128]

تجمع الخوارزمية الجينية الفوضوية (CGA) بين الخوارزمية الجينية (GA) ومشغل خرائط فوضوي (CMO) من اجل تعظيم مزايا البحث الخاصة بكلا النموذجين، اذ تستخدم الخصائص الثلاث للمتغيرات الفوضوية لتوزيع أفراد الأجيال الفرعية في المكان المعطى مما يمنع الاختيار المبكر للأجيال اللاحقة من الافراد، وتستغل الخوارزمية الجينية الفوضوية (CGA) خاصية التقارب البطيء في الخوارزمية الجينية (GA) للتغلب على عشوائية العملية الفوضوية ومن ثمّ تزيد من احتمالية انتاج افراد أفضل، وإيجاد الحل الأمثل العالمي. [125]

3.2.1.2.3: خطوات تنفيذ الخوارزمية الجينية الفوضوية المهجنة مع انحدار المتجه الداعم

يتم توضيح الخطوات الخاصة بالخوارزمية المقترحة وعلى وفق الآتي:

- الخطوة 1. توليد اعداد أولية من خلال مشغل الخرائط الفوضوي: يمكن تمثيل المعلمات

$$X_k^{(i)}, k = C, \sigma, \varepsilon \text{ (SVR)}$$

حيث ان:

(i) : تشير الى عدد التكرار

$X_k^{(i)}$: تمثل المعلمات الثلاث في نموذج انحدار المتجه الداعم (SVR), $k = C, \sigma, \varepsilon$ تعيين

$i = 0$ ، ولتعيين المعلمات الثلاث (C, σ, ε) بين المدد (Max_k, Min_k) الى متغيرات فوضوية

$x_k^{(i)}$ تقع في المدة $[0,1]$ يتم استخدام الصيغة على وفق الآتي:

$$x_k^{(i)} = \frac{X_k^{(i)} - Min_k}{Max_k - Min_k}, K = C, \sigma, \varepsilon \dots (3.7)$$

حيث ان: [70]

$x_k^{(i)}$: تمثل المتغيرات الفوضوية بعد التطبيع $X_k^{(i)}, k = C, \sigma, \varepsilon$; Min_k : تمثل الحد الأدنى

لقيم المعلمات $K = C, \sigma, \varepsilon$; Max_k : تمثل الحد الأقصى لقيم المعلمات $K = C, \sigma, \varepsilon$.

لحساب وترميز المتغير الفوضوي $x_k^{(i+1)}$ للتكرار، يتم الاعتماد على معادلة (3.8) عندما

$$\mu = 4$$

$$x_k^{(i+1)} = \mu x_k^{(i)} (1 - x_k^{(i)}) \dots (3.8), i = 0, 1, 2, 3, \dots$$

كما يتم تحويل $x_k^{(i+1)}$ للحصول على ثلاث معلمات للتكرار $X_k^{(i+1)}$ على وفق المعادلة الآتية:

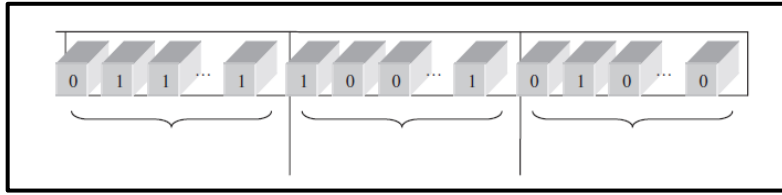
$$X_k^{(i+1)} = Min_k + X_k^{(i+1)} (Max_k - Min_k) \dots (3.9)$$

بعد التحويل، يتم ترميز المعلمات الثلاث (C, ε, σ) الى صيغة ثنائية، اذ يتم تمثيلها بواسطة

سلسلة مكونة من "جينات" من الأرقام الثنائية كما يظهر في الشكل (21)، وكل سلسلة لديها ثلاث

جينات (genes) أي تمثل ثلاث معلمات على التوالي، ويتكون كل جين من 40 (bit)، مثلاً

اذا كل (genes) يحتوي على 40 (bit)، فإن السلسلة تحتوي على 120 (bit). [52].



شكل (19) ترميز ثنائي للسلسلة. [101]

- الخطوة 2: يتم استخدام متوسط الخطأ النسبي السالب للنسبة المئوية المطلقة (MAPE-)

لتقييم اللياقة البدنية لدالة الهدف فيما يتعلق بكل حل، أي ان الخطأ في التنبؤ في نموذج

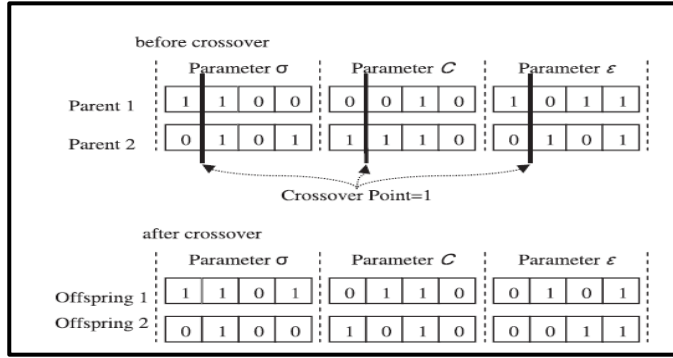
انحدار المتجه الداعم (SVR) يستخدم لتقييم اللياقة.

- الخطوة 3: من المرجح أن يتم اختيار الحلول ذات قيم اللياقة البدنية المنخفضة، اذ يتم

استخدام مبدأ عجلة الروليت لاختيار الحلول للتكرار استناداً الى دالة اللياقة البدنية. [77]

- الخطوة 4: عمليات التقاطع (Crossover). في هذه العملية يتم اختيار الحلول عشوائياً لتتطابق بعضها بعضاً. وبعد تحديد نقاط التوقف، يتم تبادل سلسلة الحلول المقترنة باستخدام مبدأ نقطة التقاطع المفردة. [52]

والشكل (20) مثال بسيط لتمثيل المعلمات (C, σ, ϵ) فنفرض ان (gene) يحتوي على أربعة (bits)، أي بمعنى ان سلسلة الحلول تحتوي على 12 (bits).



شكل (20) مثال مبسط لتمثيل المعلمات (C, ϵ, σ) . [101]

- الخطوة 5: عملية الطفرة الفوضوية. يتم تطبيع المعلمات الثلاث بعد عملية التقاطع $\hat{X}_k^{(i)}$ كمتغيرات فوضوية $\hat{x}_k^{(i)}$ $k = C, \sigma, \epsilon$ على وفق المعادلة الآتية:

$$\hat{x}_k^{(i)} = \frac{\hat{X}_k^{(i)} - Min_k}{Max_k - Min_k} \dots (3.10) \quad , k = C, \sigma, = 1, 2, \dots \dots q_{max}$$

حيث ان:

q_{max} : تمثل أقصى جيل تطوري للسكان

$\hat{x}_k^{(i)}$: تمثل المتغيرات الفوضوية بعد تطبيع $\hat{X}_k^{(i)}$, $k = C, \epsilon, \sigma$

$\hat{X}_k^{(i)}$: تمثل المعلمات الثلاث لما بعد عملية التقاطع لنموذج انحدار المتجه الداعم (SVR)

ثم يتم حساب المتغيرات الطفرية الفوضوية $\tilde{x}_k^{(i)}$ من خلال استخدام المعادلة الآتية:

$$\tilde{x}_k^{(i)} = \hat{x}_k^{(i)} + \delta x_k^{(i)} \dots (3.11)$$

حيث ان:

δ : تمثل عملية التطهير; $\tilde{x}_k^{(i)}$: تمثل متغيرات الطفرة الفوضوية $K = C, \epsilon, \sigma$.

يتم تعيين متغير الطفرة الفوضوي الذي تم الحصول عليه في المدة [0,1] الى الفاصل الزمني

للحلول (Max_k, Min_k) من خلال احتمال معين للطفرة (p_m) وعلى وفق الاتي:

$$\tilde{X}_k^{(i)} = Min_k + \tilde{x}_k^{(i)} (Max_k - Min_k) \dots (3.12)$$

حيث ان:

p_m : تمثل احتمال الطفرة.

- الخطوة 6: حالة التوقف. إذا كان عدد الحلول مساويا لمستوى معين، يتم تقديم أفضل حل،

والا يتم الرجوع الى الخطوة (2). [131]

4.2.1.2.3: تحسين معلمات انحدار المتجه الداعم باستخدام الخوارزمية الجينية الفوضوية

ان جمع القيم المناسبة لمعلمات المتجه الداعم تؤدي دوراً مهماً في اداء وكفاءة المتجه الداعم فتمثل

إحدى المشكلات المهمة في المتجه الداعم. فمن الضار اختيار المعلمات بشكل غير صحيح اذ

تعتمد دقة نموذج المتجه الداعم الى حد كبير على تحديد واختيار معلماته، لذلك تتم محاولة

لتراكيبات مختلفة للمعلمات الثلاث عن طريق اخذ عينات من مساحة البحث في فترات متقطعة.

بمجرد العثور على تركيبة مع الحد الأدنى من الخطأ التربيعي، يتم البحث عن طريق تقليل الفاصل

الزمني للعينة حول المجموعة، وتكرار العملية الى ان لا يحدث تغيير ملحوظ في دقة التحقق من

الصحة، لذلك يتم اقتراح تطبيق الخوارزمية الجينية الفوضوية (Chaotic genetic algorithm)

لتحسين معلمات المتجه الداعم. [119]

2.2.3: خوارزمية المناعة الفوضوية (chaotic Immune algorithm) واستخدامها

في تحديد معلمات انحدار المتجه الداعم (SVR)

قبل التطرق الى منهجية خوارزمية المناعة الفوضوية لابد لنا من توضيح خوارزمية المناعة لأنها

الأساس في خوارزمية المناعة الفوضوية

1.2.2.3: خوارزمية المناعة (Immune algorithm) [108]

تم اقتراح نظام خوارزمية المناعة (IA) من قبل الباحث (Mori) عام 1993 لحل مشكلات

التحسين غير الخطي، الذي يقوم على محاكاة نظام المناعة البيولوجي. فتمتع خوارزمية المناعة (IA) بالقدرة على إيجاد الحلول المثلى مع الحفاظ على درجة عالية من التنوع في مجال البحث. [130] فخوارزمية المناعة (IA) خوارزمية بحث عشوائي استكشافية لوجود قدرة تطويرية تجمع بين الاختيار والتعرف العشوائي، وقد اتفق الباحثون على ان الخوارزمية هي محاكاة أساسية لجهاز المناعة في استجابة مناعة تكيفية مع دراسة الاجسام المضادة لاستكمال العملية.

تم استخدام خوارزمية المناعة (IA) على نطاق واسع لحل مشكلات التحسين، وعلى وفق أسلوب خوارزمية المناعة (IA) تمت مقارنتها مع الخوارزمية الجينية (GA) واتضح أنها توفر طريقة اكثر كفاءة لحل مشكلات التحسين، [113] الا ان هناك عيبا واضحا في الخوارزمية (IA)، فسرعة تشغيلها بطيئة مما تؤثر سلبا في تطبيقها. تهدف خوارزمية المناعة الى تعظيم دالة اللياقة، فلدى (IA) القدرة على إيجاد افضل حل لمشكلات تحسين الأداء، عند اجراء الخوارزمية يمكن تفسير مشكلة التحسين على انها مولدات المضادات، اما الحلول الممكنة لمشكلة تحسين الأداء فيتم التعامل معها بوصفها أجساماً مضادة. [66]

1.1.2.2.3: خطوات عمل خوارزمية المناعة (Immune algorithm)

تعد خوارزمية المناعة (IA) أداة اختبار وطريقة تحسين فعالة بين الطرائق الأخرى وهي طريقة مستمدة من طريقة الأداء في الجهاز المناعي في مواجهة الامراض الخارجية مثل الفايروسات والبكتريا. وفي سياق عام من مولدات المضادات التي تستهدف الكائنات الحية من المخلوقات، اذ انها لا تدمر الخلايا فقط بل تقوم بإنتاج الخلايا أيضا. [71]

فيتم ادراج خطوات عملها على وفق الآتي:

- الخطوة 1: التهيئة العشوائية في مجموعة في مجموعة القيم الأولية (الاجسام المضادة)

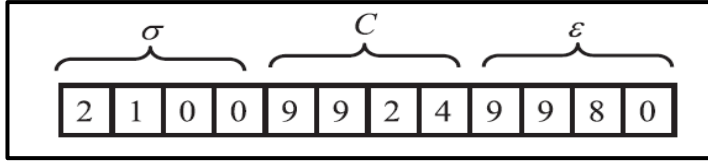
(Random initialization of antibody population): يتم توليد مجموعة من القيم

الأولية بشكل عشوائي باستخدام سلسلة الشفرة الثنائية لتمثيل المعلمات الثلاث (C, ϵ, σ)

لانحدار المتجه الداعم (SVR)، على سبيل المثال لنفرض ان القيم الاولية تحتوي على

(12) رمزاً ثنائياً تمثل ثلاث معلمات لانحدار المتجه الداعم، اذ يتم استخدام أربعة رموز ثنائية

لتمثيل كل معلمة من معلمات انحدار المتجه الداعم (SVR).



شكل (21) مثال على القيم الاولية (الاجسام المضادة). [91]

- الخطوة 2: التعرف على التقارب والتشابه (Identifying the affinity and the similarity):

(similarity) تشير قيمة التقارب الأعلى الى ان القيمة الأولية أكثر نشاطا مع مولد

المضادات، وللحفاظ على تنوع المحتفظ به في خلايا الذاكرة (the memory cells) فان

القيم الأولية ذات أوجه التشابه الأقل تملك احتمالية أكبر للاستخدام، لذا فان القيم الأولية ذات

قيمة التقارب العالية وقيمة التوافق المنخفضة لديها فرصة جيدة للدخول الى خلايا الذاكرة. فتعد

القيم الأولية الحل الأمثل لمشكلة التحسين.

يعرف التقارب بين القيم الأولية على وفق الاتي:

$$Ag_k = 1/(1 + d_k) \dots (3.13) \quad ; k = 1, 2, \dots, N$$

حيث ان:

d_k : تمثل أخطاء التنبؤ في (SVR) والتي يتم الحصول عليها من خلال k_i . [71] يتم التعبير

عن التشابه على وفق الاتي:

$$Ab_{ij} = 1/(1 + T_{ij}) \dots (3.14)$$

حيث ان:

T_{ij} : تمثل الفرق بين خطأ التنبؤ الحالي وخطأ التنبؤ السابق لنموذج (SVR) الذي تم الحصول

عليه من القيم الأولية داخل خلايا الذاكرة وخارجها. [91]

• الخطوة 3: اختيار القيم الأولية في خلية الذاكرة (Selection of antibodies in the)

(memory cell): ان القيم الأولية نوات القيم الاعلى من Ag_k تكون مرشحة جيدة لدخول خلية الذاكرة، من ناحيه أخرى، فأن مرشحي القيم الأولية المحتملون الذين يمتلكون قيم Ab_{ij} التي تتجاوز حداً معيناً غير مؤهلين للانضمام إلى خلية الذاكرة.

• الخطوة 4: اجراء عمليات التقاطع والطفرات في مجموعة القيم الأولية

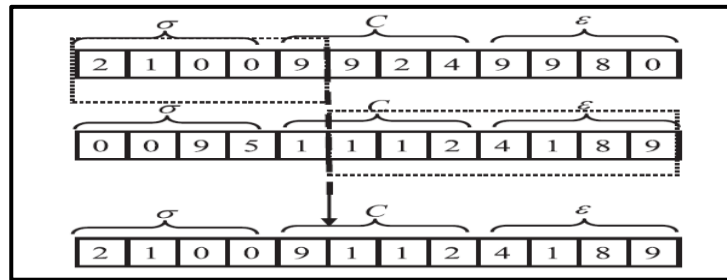
(Crossover and mutation of antibody population): تستخدم عمليات

التقاطع والطفرات لتوليد قيم أولية جديدة ، يتم اقتران الخيوط التي تمثل القيم الأولية عشوائيا

لعمل نقاط التقاطع واستخدمت عشوائيا مع نقطة التقاطع في سلسلتين من القيم الاولية اذ تم

اختيارها عشوائيا ويتم تبديل الجينات في سلسلتي القيم الأولية لإنشاء قيم أولية جديدة [71] ويمكن

تمثيل عمليات التقاطع على وفق الشكل (22).



شكل (22) عمليات التقاطع. [91]

يوضح الشكل (24) مثالا على عملية التقاطع مع قيم المعلمات الثلاث (C, ϵ, σ) لنموذج انحدار

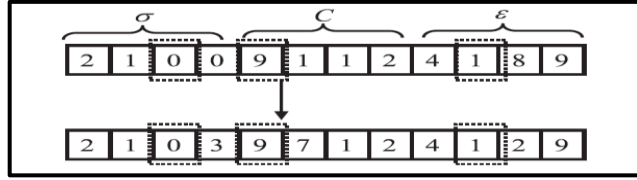
المتجه الداعم (SVR).

بعد عملية التقاطع يتم تنفيذ الطفرات بشكل عشوائي لمنع الحلول من التجانس في المستوى الامثل

المحلي مع احتمال حدوث الطفرة، من خلال تحويل الشفرة "0" الى "1" او تحويل الشفرة "1"

الى "0" كما هو موضح في الشكل (25)، وتستخدم الاحتمالات لحساب معدلات التقاطع

والطفرات. [71]



شكل (23) عمليات الطفرات. [91]

- الخطوة 5: تحديث تجمع خلايا الذاكرة (Updating the memory cell pool): القيم

الأولية ذات قيمة تقارب أعلى وقيمة تشابه اقل تكون لها فرصة جيدة للاختيار كعضو جديد

في خلايا الذاكرة للحفاظ على تنوع القيم الأولية. [91]

- الخطوة 6: معايير التوقف (Stopping criteria): اذا كان عدد الحلول يساوي حجم معين

فان القيم الأولية الأفضل توفر اصغر قيمة (MAPE) وتمثل الحل الأمثل، والا تتم العودة

الى الخطوة 2. [71]

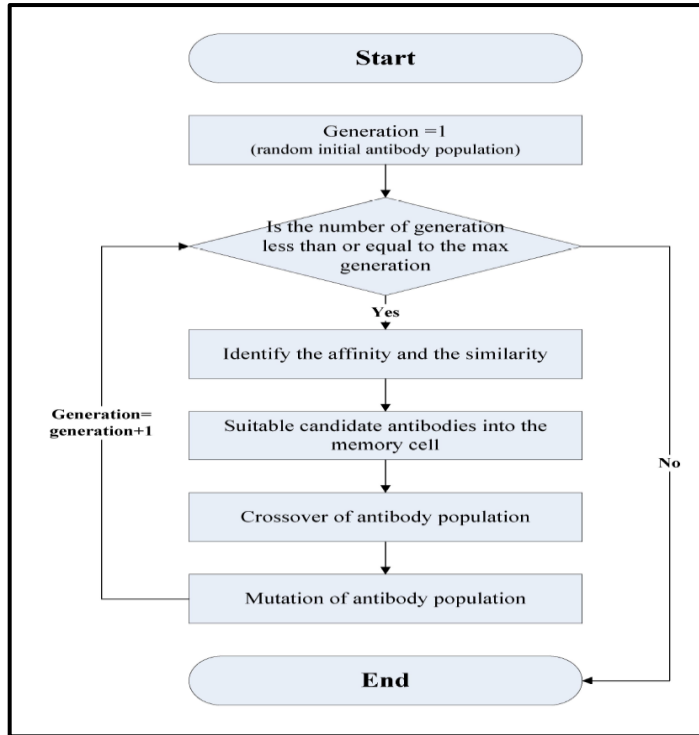
والجدير بالذكر لقد تم حل العديد من مشكلات التحسين باستخدام خوارزمية المناعة (IA) مثل

مشكلات تصنيف الأنماط، والتحسين الأمثل متعدد الأهداف، مشكلات وحدة توليد الحرارة قصيرة

الاجل، مشكلات الجدولة. وقد استخدم نظام المناعة القيم الأولية والمضادات لتمثيل الدالة

الموضوعية لخوارزمية المناعة (IA). ويتم توضيح الخطوات المذكورة آنفاً في مخطط (3)

للخوارزمية المناعية (IA).



مخطط (3) يمثل الشكل الانسيابي لخوارزمية المناعة (IA). [71].

2.1.2.2.3: ميزات خوارزمية المناعة (Immune algorithm)

تتميز خوارزمية المناعة (IA) ببعض الخصائص وعلى وفق الآتي:

1. يتم الحفاظ على التنوع بواسطة تعزيز أو تقييد أجيال الاجسام المضادة، فيتم تحقيق المستوى الأمثل للعديد من مشكلات التحسين.

2. تعمل خوارزمية المناعة (IA) على وفق الحلول الممكنة التي تضمن التقارب القوي، ومن

المتوقع الوصول الى الحل الأمثل بسرعة من خلال اجراء التحسين التكراري. [90]

الا ان خوارزمية المناعة (IA) كما هو الحال في الخوارزمية الجينية (GA) وخوارزمية تحسين

سرب الطيور (PSO) هي خوارزمية تطويرية قائمة على السكان تقدم مجموعة من الحلول

لاستكشاف واستغلال مجال البحث للحصول على الحلول المثلى. [104] على وفق إجراء خوارزمية

المناعة اذا لم يكن من الممكن الحفاظ على التنوع السكاني الاولي في ظل قيود انتقائية، فان

خوارزمية المناعة لا تستطيع البحث عن حلول الا في مساحة ضيقة والحل يكون بعيداً عن الحل

الأمثل، أي الملاءمة في الحد الأدنى المحلي. ولمعالجة هذا العيب يجب العثور على نهج فعال أو

استراتيجية او تحسينات مفيدة لخوارزمية المناعة (IA) من اجل الحفاظ على التنوع السكاني، ومراقبة مساحه الحل بدقة وكفاءه وتجنب الوصول الى المستوى الأمثل المحلي. يتلخص أحد النهج في تقسيم مجموعة السلاسل الى العديد من المجموعات الفرعية وتقييد الانتقال بين أعضاء مجموعات فرعية مختلفة، الا ان هذه الطريقة تتطلب حجما ضخما من السكان، ولضمان تنوع السكان يركز أحد هذه النهج على الفوضى وذلك نظرا لسهولة تنفيذه وقدرته الخاصة على تجنب الوقوع في المستوى الأمثل المحلي.

وفي الدراسة الحالية يتم استخدام خوارزمية المناعة الفوضوية (Chaotic Immune algorithm) لتحديد ثلاث معلمات لنموذج انحدار المتجه الداعم (Support vector regression).

2.2.2.3: خوارزمية المناعة الفوضوية (Chaotic immune algorithm)

ان خوارزمية المناعة الفوضوية (CIA) هي طريقة فعالة قابلة للتطبيق لتحسين اختيار المعلمات الثلاث في نموذج انحدار المتجه الداعم (SVR)، وقد استخدم العديد من الباحثين امثال (Wang) عام 2009 خوارزمية المناعة الفوضوية (CIA) لتقييم المعلمات الثلاث (C, ϵ, σ) لنموذج انحدار المتجه الداعم (SVR)، فقد اثبتت فعاليتها وإعطاءها أداءً جيداً في تخطي المستوى المحلي الأمثل. [72] وتجدر الإشارة الى ان دقة التنبؤات تتأثر بقيم المعلمات الثلاث (C, ϵ, σ) لنموذج (SVR) ومع ذلك لا توجد أساليب هيكلية للتحقق من كفاءه المعلمات. استخدم Hong في عام 2009 مؤخرًا خوارزمية المناعة لتحديد معلمات (SVR)، واكتشف ان هذا النموذج يتفوق على أساليب التنبؤ الأخرى مثل نماذج الشبكة العصبية (NN) ونماذج الانحدار. ان استخدام التسلسلات الفوضوية في إجراءات التحسين العشوائية، مثل الاختلافات الطفيفة في اعدادات المعلمات أو القيم الأولية في النموذج طريقة جيدة لتنوع مجال الوصف الأولي. [69]

1.2.2.2.3: خطوات عمل خوارزمية المناعة الفوضوية المهجنة مع انحدار المتجه الداعم

توجد العديد من العوامل الرئيسية المستخدمة في تصميم خوارزمية المناعة الفوضوية مطابقة مع

العوامل التي استخدمت في تصميم خوارزمية المناعة مثل اختيار القيم الأولية والتعرف على

التشابه، والتقاطع، والطفرة في القيم الأولية، وتكون خطوات عمل خوارزمية المناعة الفوضوية

(CIA) على وفق الآتي:

- الخطوة 1: تهيئة القيم الأولية (Initialization of initial values): يتم تمثيل قيم المعلمات

الثلاث في انحدار المتجه الداعم في التكرار (i) باستخدام $\sigma, C, \varepsilon, k = \sigma, C, \varepsilon, X_k^{(i)}$ ، تعيين

$i = 0$ واستخدام المعادلة (3.15) لتحويل المعلمات الثلاث بين المدد (Min_k, Max_k) الى

متغير فوضوي $x_k^{(i)}$ يقع في المدة $[0,1]$:

$$x_k^{(i)} = \frac{X_k^{(i)} - Min_k}{Max_k - Min_k} \dots (3.15)$$

ولحساب المتغير الفوضوي التالي $x_k^{(i+1)}$ يتم استخدام التسلسل الفوضوي عندما $\mu = 4$ من

خلال استخدام العلاقة التالية:

$$x^{(i+1)} = \mu x^{(i)} (1 - x^{(i)}) \dots (3.16)$$

$$x^{(i+1)} = 4x^{(i)} (1 - x^{(i)}) \quad , \quad i = 1, 2, 3, \dots, N$$

يتم تحويل $x_k^{(i+1)}$ للحصول على ثلاث معلمات للتكرار التالي $X_k^{(i+1)}$ على وفق الآتي:

$$X_k^{(i+1)} = Min_k + x^{(i+1)} (Max_k - Min_k) \dots (3.17)$$

وبعد التحويل، تشكل المعلمات الثلاث (C, σ, ε) في نموذج انحدار المتجه الداعم التعداد الاولي

للقيم الأولية والتي يتم تمثيلها بسلسلة شفرة ثنائية.

- الخطوة 2: تعريف التوافق والتشابه (Identify the affinity and the similarity):

القيم الاولية ذات قيمة تقارب أعلى لديها تنشيط مستضد أعلى. للحفاظ على مجموعة متنوعة

من القيم الأولية الموجودة في خلايا الذاكرة، من المرجح تضمين القيم الأولية ذات التشابه

الاقبل. ونتيجة لذلك، فالقيم الأولية ذات قيمة تقارب عالية ولكن قيمة تشابه منخفضة فأنها

تخترق خلايا الذاكرة.

اذ يتم تعريف التقارب بين القيم الأولية (الاجسام المضادة) على وفق المعادلة التالية:

$$Ag_k = 1/(1 + d_k) \dots (3.18)$$

كما ويتم التعبير عن التشابه بين القيم الاولية على وفق المعادلة التالية:

$$Ab_{ij} = 1/(1 + T_{ij}) \dots (3.19)$$

• الخطوة 3: اختيار القيم الأولية في خلايا الذاكرة (selection of antibodies in the

memory cell): القيم الأولية ذات القيم الأعلى لـ Ag_k لديها فرصة أفضل لدخول خلية

الذاكرة. من ناحية أخرى، لا يحق لمرشحي هذه القيم المحتملين الذين لديهم قيم تتجاوز Ab_{ij}

الانضمام إلى خلية الذاكرة.

• الخطوة 4: التقاطع للقيم الاولية (Crossover of antibody population): وتستخدم

إجراءات التقاطع والتطفير لإنتاج قيم اولية جديدة. تتم مطابقة سلاسل القيم الاولية بشكل

ديناميكي لتنفيذ عملية التقاطع. يتبع المخطط المقترح مفهوم نقطة التقاطع احادي النقطة،

عندئذ يتم تبديل السلاسل المقترنة بين نقطتي توقف محددة. [72]

• الخطوة 5: الطفرة الفوضوية في القيم الأولية (Chaotic Mutation of antibodies): يتم

تعيين مجموعة القيم الاولية في اثناء العبور ($\hat{X}_k^{(i)}, k = C, \epsilon, \sigma$) من مساحة الحل الحالية

(Min_k, Max_k) الى مدة المتغير الفوضوي [0,1] لتكوين فضاء متغير فوضوي ($\hat{x}_k^{(i)}, k =$

C, ϵ, σ) على وفق المعادلة الاتية:

$$\hat{x}_k^{(i)} = \frac{\hat{X}_k^{(i)} - Min_k}{Max_k - Min_k} \dots (3.20)$$

تتم إضافة المتغير الفوضوي $x_k^{(i)}$ الى $\hat{x}_k^{(i)}$ ، ثم يتم حساب متغير الطفرة الفوضوي في المدة

[0,1] على وفق المعادلة التالية:

$$\tilde{x}_k^{(i)} = \hat{x}_k^{(i)} + \delta x_k^{(i)} \dots (3.21)$$

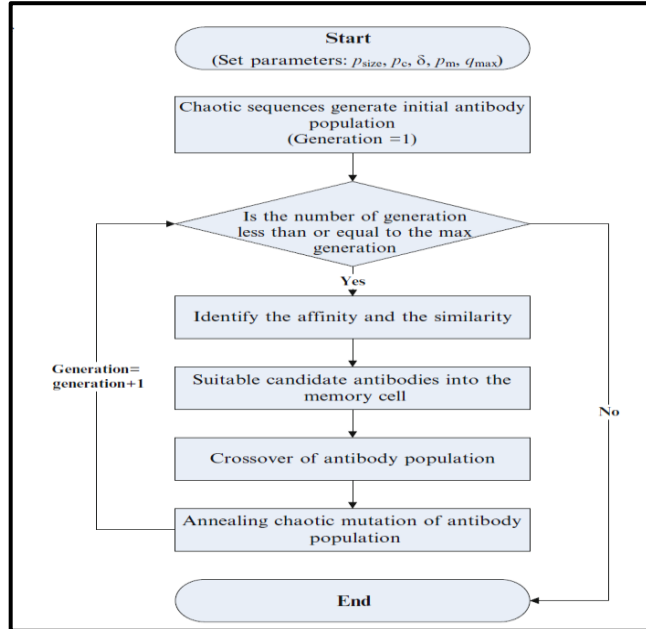
وأخيرا باستخدام احتمال حدوث الطفرة (p_m) يتم تعيين متغير الطفرة الفوضوي الذي تم الحصول

عليه في المدة [0,1] الى مدة الحل (Min_k, Max_k) على وفق المعادلة الاتية:

$$\tilde{X}_k^{(i)} = Min_k + \tilde{x}_k^{(i)}(Max_k - Min_k) \dots (3.22)$$

- الخطوة 6: معايير التوقف (Stopping criteria): عندما يكون عدد الحلول يساوي مستوى معيناً فإن القيم الأولية الأفضل تمثل الحل، وألا تتم العودة الى الخطوة الثانية. [69] ويوضح

مخطط (4) إجراءات عمل خوارزمية المناعة الفوضوية (CIA)



مخطط (4) خوارزمية المناعة الفوضوية (CIA). [69]

3.2.3: خوارزمية تحسين سرب الطيور (particle swarm optimization algorithm)

واستخدامها في تحديد معاملات نموذج انحدار المتجه الداعم (SVR)

1.3.2.3: خوارزمية تحسين سرب الطيور (PSO)

تعد من الخوارزميات التي تحاكي سلوك المجموعات مثل مجموعات الطيور، الأسماك، الحشرات

وما الى ذلك. تنتمي خوارزمية تحسين سرب الطيور (PSO) الى فئة تقنيات نكاء السرب

المستخدمة في حل مسائل الامثلية وقد طبقت هذه الخوارزمية عام 1995 من قبل الباحثين

(Kennedy) و (Eberhard)، وتعتمد خوارزمية (PSO) على مجموعة من القيم العشوائية

تسمى الجسيمات (particle) وكل جسيم يمتلك سرعة وموقعاً خاصاً يمكن تغييره، ومن ثمَّ

تتحرك هذه الجسيمات في مساحة بحث عشوائية، وتدعى هذه الخوارزمية أيضاً بخوارزمية سرب

الجسيمات. [30]

ان الهدف من خوارزمية تحسين سرب الطيور (PSO) الحصول على الحل الأمثل والأفضل من

خلال محاكاة سلوكيات الطيور عند البحث عن الطعام الأفضل لذلك فإن أسلوب هذه الخوارزمية سيتم تشكيلا في البداية بعملية تهيئة عشوائية لمجموعة من السكان وعدة حلول مرشحة في فضاء بحث المشكلة، ويتم البحث ضمن هذه المجموعة عن الحل الأمثل. [8]

تمثل خوارزمية (PSO) تقنية تحسين قائمة على السكان اقترحت لأول مرة لحل مشكلة التقليل الى أدنى حد غير مقيد، اذ يمكن استخدام مشكلة التحسين الوظيفي في صياغة العديد من مشكلات التحسين الحقيقي وكالاتي:

$$\min \quad f(X), X = [x_1, x_2, \dots, x_n] \dots (3.23)$$

$$S.t \quad x_i \in [a_i, b_i], i = 1, 2, 3, \dots, n$$

حيث ان:

f : تمثل الدالة الموضوعية ; X : يمثل متجه القرار المكون من n من المتغيرات. [103; p.1262]

يغير الجسيم سرعته وموقعه في كل تكرار من خلال تعقب (P_{ibest}) الحد الأقصى الفردي (الحل الأمثل للجسيم) والمعدل الأقصى العالمي (g_{best}) (الحل الأمثل الذي يجده السرب بأكمله). لذا تمثل تقنية الحوسبة التطورية القائمة على التحسين التكراري، فتستخدم التكرارات لإيجاد أفضل حل لمجموعة من الجسيمات العشوائية، فتعد خوارزمية ذكية للتحسين، اذ يمكن وصف العملية على

النحو الاتي:

في فضاء بحث متعدد الأبعاد (d – dimensional)، يفترض ان هناك جسيمات (M) لتكوين

سرب. إذ يتم التعبير عن الجسيم (i) على انه متجه عالي الابعاد (x_i)

$$x_i = (x_{i1}, x_{i2}, \dots, x_{id})^T \dots (3.24)$$

أي ان موقع الجسيم (i) في فضاء البحث ذي الابعاد العالية ، هو x_i كمتجه عالي الابعاد وتمثل سرعة طيران الجسيم i على النحو الاتي:

$$v_i = (v_{i1}, v_{i2}, \dots, v_{id})^T \dots (3.25)$$

فالموقع الأفضل للجسيم i يكون على وفق الاتي:

$$p_i = (p_{i1}, p_{i2}, \dots, p_{id}) \dots (3.26)$$

$$p_g = (p_{g1}, p_{g2}, \dots, p_{gd}) \dots (3.27)$$

تستخدم المعادلات الاتية في خوارزمية تحسين سرب الطيور (PSO) لتغيير سرعة الجسيم وموقعه.

$$v_{id}^{(k+1)} = \left(w^{k+1} v_{id}^{(k)} + c_1 r_1 (p_{id}^{(k)} - x_{id}^{(k)}) + c_2 r_2 (p_{gd}^{(k)} - x_{id}^{(k)}) \right) \dots (3.28)$$

يتم تحديد الموقع الجديد للجسيم لكل معلمة في الجيل التالي على وفق المعادلة الآتية: [115]

$$x_{id}^{(k+1)} = x_{id}^{(k)} + v_{id}^{(k)} \dots (3.29)$$

$$(i = 1, 2, \dots, N) ; (d = 1, 2, \dots, D) ; k = c, \varepsilon, \sigma$$

حيث ان:

w : هو معامل وزن القصور الذاتي، ويعبر عنه على وفق الآتي:

$$w^{(k)} = \frac{t}{T_{max}} (w_{fin} - w_{ini}) + w_{ini} \dots (3.30)$$

w_{ini} : تمثل الوزن الاولي ; w_{fin} : تمثل الوزن النهائي ; t : تمثل التكرار الحالي ; T_{max} : تمثل

اقصى عدد من التكرارات ; c_1, c_2 : تمثل ثوابت اختيارية غير سالبة ; r_1, r_2 : تمثلان دالتين

عشوائيتين تكون بين (0,1). [94]

$x_i = (x_{i1}, x_{i2}, \dots, x_{id})$: تمثل الجسيمات.

$p_i = (p_{i1}, p_{i2}, \dots, p_{id})$: يمثل الموقع لأفضل قيمة لياقة للجسيمات.

$p_g = (p_{g1}, p_{g2}, \dots, p_{gd})$: تمثل افضل جسيم بين الجسيمات في المجتمع.

$v_i = (v_{i1}, v_{i2}, \dots, v_{id})$: تمثل معدل تغير سرعة الجسيم.

D : يمثل بعد الجسيم، h : يمثل عدد الجسيم.

اذ توضح المعادلة (3.28) كيفية التحديث الديناميكي لسرعة الجسيمات (الطيور). اما المعادلة

(3.29) فهي تصف كيفية تحديث الموقع للجسيمات. [36]

تكون خوارزمية (PSO) بسيطة في المفهوم وسهلة التنفيذ وعلى الرغم من ان معادلاتها لا تحتوي

على أي تعقيد الا انها تتمتع بكفاءة حسابية عالية، وقد لاقت الكثير من الاهتمام والتطبيقات

الواسعة في حل مشكلات التحسين غير الخطي بسبب سهولة تنفيذها والمفهوم البسيط والتقارب

السريع. [49]

1.1.3.2.3: خطوات عمل خوارزمية تحسين سرب الطيور (PSO)

تشتمل خوارزمية (PSO) على توليد اعداد عشوائية إذ يتأثر أداؤها بكمية ونوعية الأرقام المولدة،

ويمكن ادراج خطوات عملها على وفق الاتي:

1. تهيئة موقع وسرعة السرب: تحديد مجموعة من الجسيمات ذات موقع عشوائي وسرعة ابتدائية

(أولية) في منطقة البحث.

2. تقييم لياقة الجسيمات: تقييم قيمة اللياقة لكل جسيم (حل).

3. مقارنة لياقة كل جسيم مع p_{best} (personal best) : يتم تعيين قيمة p_{best} مساوية للقيمة

الحالية والموقع p_i مساويا للموقع الحالي x_i في فضاء متعدد الابعاد، اذا كانت القيمة

الحالية افضل من p_{best} .

4. المقارنة مع g_{best} (global best): مقارنة اللياقة البدنية مع افضل الخصائص العامة

السابقة. إذا كانت القيمة الحالية اكبر من g_{best} ، تتم إعادة تعيين g_{best} الى قيمة الجسيم

الحالي.

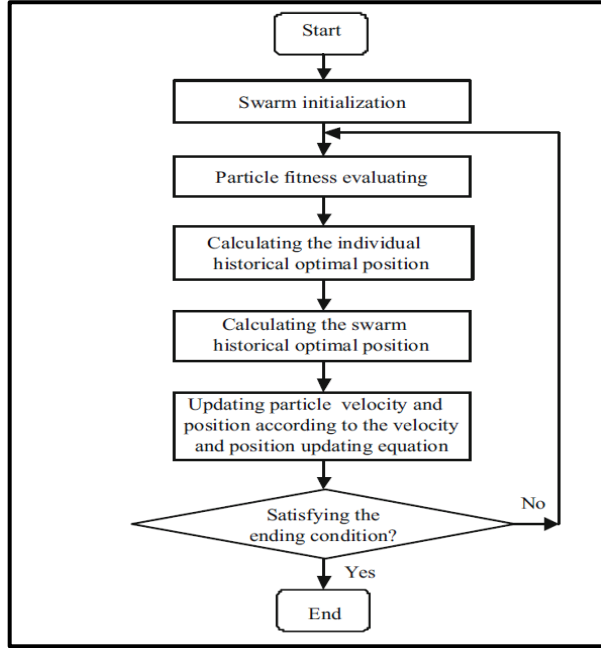
5. تحديث سرعة كل جسيم وموقعه: اذ يتم تغيير سرعة v_i وموقع x_i لكل الجسيمات على وفق

المعادلة (3.28) و (3.29).

6. تكرار الدورة التطورية: العودة الى الخطوة (2) ليتم استيفاء معيار التوقف، وعادة ما يمكن

الوصول الى الحد الأقصى لعدد التكرارات او تكون قيمة دالة اللياقة جيدة بما يكفي [50]

[36]، ويوضح المخطط (5) خطوات عمل هذه الخوارزمية.



مخطط (5) خوارزمية تحسين سرب الطيور (PSO). [115]

اهتم العديد من الباحثين في مجال تحسين الأداء بهذه الخوارزمية لما لها من دور كبير في التطبيقات الفعالة اذ شملت شتى المجالات فهي مثالية لحل مسائل الامثلية غير الخطية، ولا تزال

خوارزمية تحسين سرب الطيور (PSO) اداة فعالة جدا في شكلها الحالي. [129]

تعمل خوارزمية سرب الطيور (PSO) بإبقاء عدة حلول مرشحة في فضاء البحث، اذ يتم تقييم كل

حل مرشح من خلال دالة الهدف التي تم تحسينها خلال كل تكرار من الخوارزمية وتقييم لياقتها.

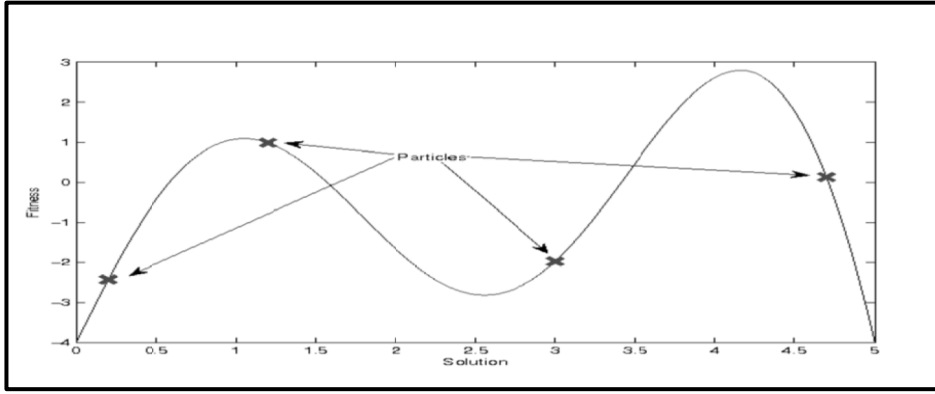
يمكن اعتبار كل حل مرشح على انه جسيم خلال مرحلة اللياقة البدنية بحثا عن الحد الأقصى او

الحد الأدنى لدالة الهدف. يتم اختيار الحلول المرشحة عشوائيا من فضاء البحث بشكل عشوائي

باستخدام دالة الهدف ثم تعمل على قيم اللياقة الناتجة. اذ يظهر الشكل (24) الحالة الأولية

للخوارزمية المكونة من أربعة جسيمات التي تجد الحد الأقصى في فضاء بحث احادي البعد.

ويتكون فضاء البحث من جميع الحلول المحتملة على طول المحور السيني.



الشكل (24) حالة تحسين سرب الطيور (PSO) الأولية. [35]

ومن الجدير بالذكر ان خوارزمية (PSO) ليس لديها أي فكره عن دالة الهدف الأساسية لذا ليس

لها طريقة لمعرفة ما اذا كانت جميع الحلول المرشحة متشابهة او بعيدة عن الحد الأقصى

المحلي. [35]

2.1.3.2.3: معامل وزن القصور الذاتي (Adaptive inertia weight factor)

يستخدم (AIWF) للسماح للجسيمات الجيدة (أزواج) بمراجعة اكتشافها من أجل تحسين النتائج

باستخدام البحث المحلي، في حين يتم تشجيع الجسيمات الفقيرة على تغيير مساحة المسح

باستخدام خطوات كبيرة. ويمثل المعدل الذي يتحكم في تأثير السرعة السابقة على المعدل الحالي،

ومن ثمَّ فان التحكم السليم في وزن القصور الذاتي امر بالغ الأهمية لإيجاد الحل الأمثل والأكثر

دقة وكفاءة. فقد نجح (Shi and Eberhart) في تحسين كفاءه (PSO) والى حد كبير بفضل

وزن القصور الذاتي (AIWF) المتغير خطيا. وتكون صيغة (AIWF) على وفق الاتي:

$$l = \begin{cases} l_{min} + \frac{(l_{max} - l_{min})(f_i - f_{min})}{f_{avg} - f_{min}}, & f_i \leq f_{avg} \\ l_{max} , & f_i > f_{avg} \end{cases} \dots (3.31)$$

حيث ان:

l_{min} : يمثل الحد الأدنى لـ (l) ; l_{max} : يمثل الحد الأعلى لـ (l) ; f_i : تمثل القيمة الموضوعية

الحالية لزوج الجسيمات.

تمثل القيم الموضوعية المتوسطة والدنيا لكل ازواج الجسيم، على التوالي.

تتفاوت قيمة (l) حسب القيم الموضوعية للجسيم اذ بالإمكان الحفاظ على الجسيمات ذات القيم الموضوعية المنخفضة في حين يتم تعطيل الجسيمات ذات القيم الموضوعية فوق المتوسط وهذا يعني ان الجسيمات الجيدة من المرجح ان تستخدم البحث المحلي لتحسين أدائها، في حين ان الجسيمات السيئة تميل الى اجراء تعديلات كبيرة لاستكشاف الفضاء بخطوة كبيرة. بعبارة أخرى يشكل وزن القصور الذاتي (AIWF) وسيلة جيدة لضمان التنوع السكاني مع الحفاظ على إمكانات تقارب جيدة. [92]

3.1.3.2.3: مزايا خوارزمية تحسين سرب الطيور (Partial Swarm Optimatizion)

تعد هذه الخوارزمية واحدة من أفضل الخوارزميات المستخدمة في حل تقنيات متعددة، اذ تمتاز ببعض المزايا التي تدرج على وفق الآتي:

1. تعد خوارزمية بحث فعالة جدا وخالية من المشتقات.
2. ان تحسين السرب هو شكل من اشكال التحسين يركز على الذكاء.
3. لا تحتوي عملية تحسين سرب الجسيمات (PSO)، على حسابات متداخلة اذ تستخدم سرعة الجسيمات في اجراء الاختبار، وسيشارك الجسيم الأكثر إيجابية المعلومات مع الجسيمات الأخرى وخلال تطور أجيال متعاقبة لذا فان سرعة الاستكشاف والبحث تكون سريعة جدا.
4. ان استخدام سرب الجسيمات للتحسين الأمثل أساسي وبسيط جدا، فهي تتمتع بإمكانية تحسين أفضل ويمكن تنفيذه بسرعة أكبر من العمليات الحسابية الأخرى.
5. تعتمد على نقاط أولية اقل مقارنة مع طرائق التحسين الأخرى.
6. لها عدد قليل من المعلمات، كما ان تأثير المعلمات في الحلول ضئيل بالمقارنة مع

استراتيجيات التحسين الأخرى. [39]

الا انه يعاب عليها في نقاط أخرى هي:

- 1- تعتمد هذه الطريقة على الثقة الجزئية، وهذا يؤدي الى انخفاض الدقة في التحكم بالسرعة والتوجيه.
 - 2- لا توجد طريقة لنشر هذا النهج وتحسينه.
 - 3- عندما لا يكون هناك نظام إحدائيات كحل في مجال الطاقة وأي حركة لقواعد الجسيمات في مجال الطاقة، فإن هذه التقنية لا يكون لها حل.
- تعتمد خوارزمية تحسين سرب الطيور (PSO) على ذكاء سرب الجسيمات على النقيض من الخوارزميات الأخرى لا تزال قائمة على أساس رياضي وحسابات منهجية. [113]
- يمكن تطبيق خوارزمية (PSO) في مجالات عدة أهمها كالاتي: [35]
1. تدريب الشبكة العصبية (Training of neural network)
 2. تحسين البنية (Structural optimization)
 3. معالجة الصور (Images processing)
 4. تحسين شبكة توزيع الطاقة الكهربائية (optimization of electric power distribution network)
 5. معالجة الكيمياء الحيوية (Process biochemistry)
 6. تعريف النظام في الميكانيكا الحيوية (System identification in biomechanics)
- يكون من السهل استخدام خوارزمية تحسين سرب الطيور (PSO) نتيجة لمبدئها البسيط وقلّة معلماتها ونتائجها الواضحة التنفيذ. فتستخدم غالبا لحل مشكلات التحسين غير الخطي ودوال متعددة غير مميزة ومعقدة. وتمثل تقنية الحل التي تبدأ من حلول عشوائية من خلال الحسابات التكرارية لإيجاد الحل الأمثل، وبعد ذلك تستخدم اللياقة البدنية لتقييم جودة الحل. ويمكن تحسين هذه الخوارزمية لتلبية احتياجات التطبيقات العملية مما يثبت إمكانية تطبيقها. [110]

فالتحكم في الاستكشاف العالمي والاستغلال المحلي أمرٌ بالغ الأهمية في خوارزمية تحسين سرب الطيور (PSO) لاكتشاف أفضل الحلول بكفاءة. تتأثر كفاءة (PSO) بوضوح بمعلماتها. يمثل الجزء الأول من المعادلة (3.28) تأثير السرعة السابقة، والتي توفر الزخم المطلوب للجسيمات التي تتحرك حول مساحة البحث. فتمثل w المعامل الذي يتحكم في تأثير السرعة السابقة في السرعة الحالية. ونتيجة لذلك، تحدد w أهمية التوازن بين الاستكشاف والاستغلال ومن ثمَّ فإن التحكم السليم في وزن القصور الذاتي أمر بالغ الأهمية لتحديد الحل الأفضل بشكل فعال وموثوق. [92] ولمعالجة هذا النقص استخدم الباحثون أسلوب خوارزمية تحسين سرب الطيور الفوضوي (CPSO) التي تجمع بين خوارزمية تحسين سرب الطيور (PSO) وعامل وزن القصور الذاتي (AIWF) والبحث المحلي الفوضوي (CLS). [43]

2.3.2.3: خوارزمية تحسين سرب الطيور الفوضوية (CPSO) [92]

استناداً إلى خوارزمية (PSO) المقترحة مع (AIWF) والبحث المحلي الفوضوي (CLS) يتم اقتراح نهج تكراري على مرحلتين يسمى (CPSO)، إذ يتم استخدام (PSO) مع (AIWF) للاستكشاف العالمي من خلال تحديث سرب الجسيمات إذ يتم استخدام (CLS) للبحث الموجه محلياً لحلول (PSO)، كما يتم استخدام ديناميكيات الفوضى لتعديل أفضل الجسيمات المولدة بواسطة (PSO) محلياً. علاوة على ذلك، يتم توليد الجسيمات الجديدة عشوائياً وإدخالها مع السكان الجدد لضمان التنوع السكاني. ولتسريع التقارب، يتم توليد جسيمات جديدة ديناميكياً.

تقوم خوارزمية تحسين سرب الطيور الفوضوية (CPSO) بتحسين خوارزمية سرب الطيور من خلال دمجها مع نظرية الفوضى، ولتوليد متغيرات فوضوية تستخدم المعادلة اللوجستية الآتية:

$$x_{n+1} = \mu x_n (1 - x_n) \dots (3.32)$$

1.2.3.2.3: البحث المحلي الفوضوي (chaotic local search)

يتم دمج الديناميكيات الفوضوية في (PSO) مع (AIWF) من أجل إثراء سلوك البحث ومنعها من الوقوع في المستوى المحلي الأمثل

وفيما يلي خطوات اجراء البحث المحلي الفوضوي (CLS) وعلى وفق الاتي:

- الخطوة 1: تعيين $(i = 0)$ واستخدام المعادلة (3.33) لتحويل المعلمات الثلاث

الى متغير فوضوي $x_k^{(i)}$ ($k = c, \sigma, \varepsilon$) من الفترات $(x_{\min(k)}, x_{\max(k)})$ تقع

في المدة الزمنية $[0,1]$. [72].

$$x_k^{(i)} = \frac{X_k^{(i)} - x_{\min(k)}}{x_{\max(k)} - x_{\min(k)}} \dots (3.33)$$

- الخطوة 2: حساب المتغير الفوضوي $x_k^{(i+1)}$ ، عندما $\mu = 4$ باستخدام العلاقة الاتية:

$$x_k^{(i+1)} = \mu x_k^{(i)} (1 - x_k^{(i)}) \dots (3.34)$$

$$x_k^{(i+1)} = 4x_k^{(i)} (1 - x_k^{(i)}) \quad ; \quad k = c, \sigma, \varepsilon ; i = 1, 2, 3, \dots, N$$

- الخطوة 3: الحصول على ثلاث معلمات للتكرار التالي $X_k^{(i+1)}$ وتحويل $x_k^{(i+1)}$ من خلال

استخدام المعادلة الاتية: [49].

$$X_k^{(i+1)} = x_{\min(k)} + x_k^{(i+1)} (x_{\max(k)} - x_{\min(k)}) \dots (3.35)$$

- الخطوة 4: حساب القيمة الموضوعية الجديدة مع $X_k^{(i+1)}$

- الخطوة 5: ان المتغير الفوضوي الجديد $x_k^{(i+1)}$ وقيمته الموضوعية المقابلة هي الحل النهائي

اذا تم الوصول الى القيمة الموضوعية الجديدة ذات قيمة مؤشر دقة التنبؤ الأصغر او الحد

الأقصى للتكرار، والا يتم تعيين $i = i + 1$ والعودة الى الخطوة 2.

في أثناء كل تكرار للتنبؤ يتم استخدام خوارزمية تحسين سرب الطيور الفوضوية (CPSO) للبحث

عن مزيج افضل من المعلمات الثلاث لانحدار المتجه الداعم (SVR). [72].

2.2.3.2.3: خطوات عملية خوارزمية تحسين سرب الطيور الفوضوية (CPSO) المهجنة مع

انحدار المتجه الداعم (SVR)

تطبق خوارزمية تحسين السرب الفوضوي (CPSO) لاختيار امكانية البحث في ثلاث معلمات في

نموذج انحدار المتجه الداعم (SVR)، اذ يستخدم عامل وزن القصور الذاتي (AIWF) والبحث

المحلي الفوضوي (CLS) للسماح للجسيمات الجيدة (الأزواج) بمراجعة اكتشافها من اجل تحسين

النتائج باستخدام خطوات اكبر للبحث المحلي والجسيمات الفقيرة لتعديل مساحة البحث.

يمكن ادراج الخطوات الأساسية لعمل هذه الخوارزمية (CPSO) وعلى وفق الاتي:

1. تهيئة أوضاع السرب (Initialization): تحديد مجموعة معينة من ازواج الجسيمات

$(C_i, \varepsilon_i, \sigma_i)$ ذات مواضع عشوائية $(X_{ci}, X_{\varepsilon i}, X_{\sigma i})$ وسرعة ابتدائية (أولية)

$(V_{ci}, V_{\varepsilon i}, V_{\sigma i})$ في منطقة البحث إذ يحتوي كل جسيم على (n) من المتغيرات.

2. حساب القيمة الموضوعية (Objective value computing): حساب القيم الموضوعية

لأزواج الجسيمات، والسماح لأفضل وضع لكل زوج من الجسيمات $(P_{ci}, P_{\varepsilon i}, P_{\sigma i})$ والقيمة

الموضوعية f_{best} لتكون مساوية لموقعها الاولي وقيمتها الموضوعية. وليكن افضل وضع

لزواج الجسيمات الاولي وقيمتها الموضوعية تساوي افضل وضع عالمي $(P_{cg}, P_{\varepsilon g}, P_{\sigma g})$

وقيمتها الموضوعية $f_{globalbest}$. [64]

3. تقييم القيمة الموضوعية (Evaluation of the objective values): تقييم القيم

الموضوعية على وفق المعادلة (3.28) و (3.29) لكل جسيم وتحديث السرعة والموضع لكل

زوج من الجسيمات.

4. المقارنة والتحديث (Comparison and update): تتم مقارنة القيمة الموضوعية الحالية

لكل زوج من الجسيمات مع f_{best} ، فاذا كانت القيمة الموضوعية الجديدة اعلى (أي أن قيمة

مؤشر دقة التنبؤ أصغر)، عندئذ يتم استبدال الوضع الأمثل $(P_{ci}, P_{\varepsilon i}, P_{\sigma i})$ وقيمتها

الموضوعية بالموضع الحالي.

5. تحديد افضل زوج من الجسيمات (The best particle pair determination): استنادا

إلى أعلى قيمة موضوعية، يتم تحديد أفضل زوج من الجسيمات. يتم استبدال $(P_{cg}, P_{\varepsilon g}, P_{\sigma g})$

وقيمتها الموضوعية مع افضل زوج من الجسيمات الجديد اذا كانت القيمة الموضوعية اقل من

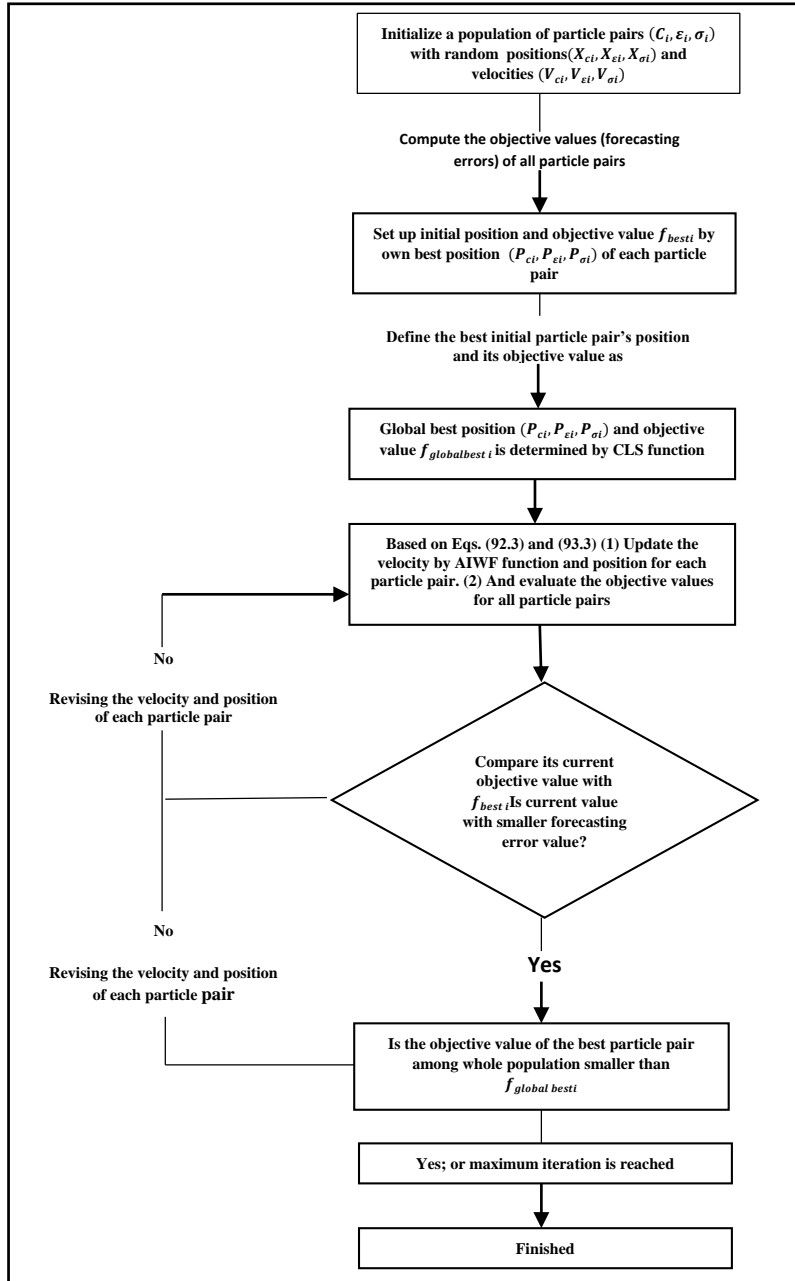
$f_{globalbest}$

6. معيار التوقف (Stop criterion): تتم العودة الى الخطوة (3) ليتم الوصول الى دقة

التنبؤ، عندها تحدد $(P_{cg}, P_{eg}, P_{\sigma g})$ و $f_{globalbest}$ ، وتكون قيمة دالة اللياقة جيدة بما يكفي

أو يمكن الوصول الى الحد الأقصى لعدد التكرارات (أجيال). [134], [72] ويبين الشكل الآتي

خطوات عمل خوارزمية تحسين سرب الطيور الفوضوية (CPS)



مخطط (6) عمل خوارزمية تحسين سرب الطيور الفوضوية (CPSO). [65]

3.2.3.2.3: تحديد دالة اللياقة البدنية (the fitness function) [94]

يتم اختيار دالة اللياقة البدنية لتمثل مجموع مربعات أخطاء التقدير.

نفترض ان السلسلة الزمنية هي x_N وقد نتج عن نموذج انحدار المتجه الداعم (SVR) تنبؤات وهي y_1 و y_2 ، فالتنبؤ المشترك هو $(\omega_1 y_1 + \omega_2 y_2)$ ، يتم التعبير عن الانحرافات بين التنبؤ والقيم الحقيقية على وفق الاتي:

$$z_i = \omega_1 y_1(i) + \omega_2 y_2(i) \dots (3.36) \quad ; \quad i = 1, 2, 3, \dots, N$$

يتم تقدير دالة لياقة خوارزمية تحسين سرب الطيور الفوضوية (CPSO) باستخدام المعادلة الآتية:

$$f = \min \sigma^2 = \min \sum_{i=1}^N (\omega_1 y_1(i) + \omega_2 y_2(i) - x_i)^2 \dots (3.37)$$

$$s. t \quad \omega_1 + \omega_2 = 1 \quad ; \quad \omega_1 \geq 0, \omega_2 \geq 0$$

4.2.3.2.3: تحسين معاملات انحدار المتجه الداعم (SVR) استنادا الى خوارزمية تحسين

سرب الطيور الفوضوية (CPSO)

تعد المعلمة غير الحساسة ومعلمة التنظيم ومعلمة دالة النواة اهم معاملات انحدار المتجه الداعم. اذ ان مدى هذه المعلمات يكون:

$$\varepsilon \in [0, 0.07] \quad , \quad C \in [1, 10] \quad , \quad \sigma \in [0, 2]$$

الخطوات الرئيسية لتحسين معاملات انحدار المتجه الداعم باستخدام (CPSO) يمكن تلخيصها على وفق الاتي:

1- انشاء متجه ثلاثي الابعاد (three – dimensional) على وفق النظام الفوضوي، ووزن

القصور الذاتي، وتهيئه الحد الأقصى للتكرار، وحد الخطأ للياقة البدنية، وحد سرعة كل جسيم.

2- تحديث أفضل حل مثالي للجسيم وللمجموعة بأكملها.

3- اجراء تحسين فوضوي من خلال تعيين (50%) من الجسيمات الأفضل أداء الى المدة

$$[0, 1].$$

4- إذا كان الجسيم المحسن من قبل نظرية الفوضى يتميز بمستوى اعلى من اللياقة فيجب تحديث أفضل جسيم.

5- يتم تكرار الخطوة (3) اذا لم يصل الى الحد الأقصى للتكرارات. [89]

3.3: النمذجة والتقلبات والتنبؤ (Volatility & forecasting, modeling)

تعد النمذجة والتقلبات والتنبؤ خطوه مهمة في السلسلة الزمنية للأحمال الكهربائية.

1.3.3: النمذجة

هي المرحلة الأولى لحل المشكلة، في هذه المرحلة يتم بناء نموذج نظام يحاكي النموذج الحقيقي

بحيث تكون خصائصه تمثل خصائص النظام الحقيقي، بعد ذلك يتم تحديد خصائص النظام

بحيث يتكامل النموذج مع النموذج الحقيقي من حيث التكلفة وسهولة الاستخدام.

فالنمذجة هي مجموعة من العمليات والمعالجات لبناء النموذج وهي قدرة النموذج على تعريف

المشكلة ووصفها بالشكل الذي يجعلها مبسطة والتي يراد بها تسهيل الظاهرة المعقدة.

يتم استخدام النمذجة لتوضيح العمليات والظواهر وخاصة العمليات والظواهر المعقدة وتبسيط

صورها، مما يساعد على تضيق الفجوة بين الواقع والنظرية. [10] فتعرف النمذجة بأنها أسلوب

رياضي تستخدم لاكتشاف الميزات والصفات والسلوك والقدرات والخصائص وترميزها وإعادة

صياغتها بطريقة بسيطة وذات مغزى، أي تمثل القدرة على تكرار الأشياء أو نقلها أو اظهارها

ولكن بشكل يسهل فهمها، أي اكتشاف خصائص وصفات الظاهرة المدروسة وتلخيصها ونقلها

للتعبير عن الظاهرة. [13]

وقد عرفها (الخفاجي، 2016) بأنها عملية تبسيط لظاهرة أو حدث نرغب في فهمها بطريقة

أفضل. اما (قرني، 2013) فقد عرفها بأنها عباره عن نظام من الأفكار تستعمل في تكوين خطة

أو آلية لتوضيح وتفسير كيفية عمل وحدث بعض العمليات والاحداث والظواهر بما يساعد في

التنبؤ بإمكانية حدوثها. [16]

فالنمذجة هي عملية تحول المشكلة الحياتية الى مسالة رياضية، ثم معالجة المشكلة وحلها واختيار

أفضل الحلول التي تتناسب مع طبيعة المشكلة التي نعالجها والتي يسهل التعامل معها ومن ثم

التعميم والتنبؤ بها. [14] وتعتمد عملية النمذجة على ما يعرف بمعالجة المعلومات والتي تتمثل

بالعمليات العقلية الانتقائية لبناء أنواع مختلفة من المعلومات والأفكار والمعرفة والثقافة. فعملية النمذجة تتطلب فهم خصائص بيانات عينة البحث من اجل وصف النموذج بشكل صحيح وتجنب أخطاء القياس والتقدير. [1]

1.1.3.3- أهمية النمذجة (Importance of modeling) [20] ، [13]

تدرج أهمية النمذجة على وفق الاتي:

1. تسهيل حل المشكلات اذ ان عرض المشكلات بصورتها الواقعية او القريبة من الواقع أفضل من عرضها بشكل مصطنع، من خلال وصف وتمثيل الظاهرة التي تم البحث عنها والتعبير عنها، وبذلك يكون لدينا فهم أعمق للظواهر والاحداث المعقدة.
2. تؤدي النمذجة دوراً حيوياً في اكتساب المعرفة الأساسية والمهمة في الإحصاء والمجالات الأخرى، اذ تعمل على تطوير عمليات التفكير المنطقي وتدريب الباحثين على صياغة اسئلة استقصائية توجههم نحو التوصل الى الحلول البحثية للمشكلات مع استخدام الأدلة والحجج المنطقية لمشكلاتهم.
3. اكتساب أسلوب علمي لحل المشكلات وتعلم صيغ جديدة من السلوك الخاص بالعمل والمهارات والقيم والعادات لأنها تعد من الطرائق المناسبة للباحثين لتطوير الاستيعاب العام من خلال ممارسة الاستدلال في اثناء تمثيل وتطبيق النموذج.
4. تساعد النمذجة على توظيف تكنولوجيا الحاسبات، اذ تسهل تقديم الاستجابة الى العمل والتوجهات والتعليمات والارشادات، ويساعد ظهور البرامج الإحصائية الهائلة على استخدام النمذجة للتنبؤ بالمستقبل.
5. تدريب الطلبة وتشجيعهم على استخدام المكتبة والحصول على المعلومات من مصادرها الأساسية.

2.1.3.3- مراحل النمذجة (Modeling stages)[13]

يتم تحديد عدة مراحل للنمذجة وعلى وفق الآتي:

1. تحديد المشكلة المراد دراستها في الواقع، ثم التعرف على العوامل والمتغيرات المؤثرة بها.
2. تحديد العلاقة بين المتغيرات وتعريف المصطلحات غير الواضحة في المشكلة.
3. محاولة صياغة نموذج باستخدام العوامل المتعلقة بالمشكلة اذ يتم ترميز كل متغير برمز خاص به.
4. التأكد من الصياغة بحيث يتطابق الهدف مع النموذج.
5. اختبار محاكاة النموذج للواقع وذلك باختبار قدرته على تمثيل الواقع وذلك بإعطاء بعض المتغيرات قيما إضافية واقعية ومقارنة نتائج المشاهدة مع نتائج النموذج لاختبار قدرة النموذج على التنبؤ بما سيحدث مستقبلا.
6. تطوير النموذج بحيث يؤدي استخدام النموذج الى القاء الضوء على المزيد من المتغيرات الأقل أهمية بهدف دمجها في النموذج الأصلي للحصول على نموذج معدل يعطي نتائج أفضل، وإمكانية توسيع مجال استخدامه بحيث يصلح لمواقف جديدة من خلال دراسة علاقة النموذج بنماذج أخرى في المجال نفسه.
7. تفسير واختبار نتائج النموذج وملاحظة سلوك النموذج والحصول على الحل الأمثل. والعودة للواقع ومحاولة التأكد من إمكانية تطبيق الحل على الواقع لمعرفة فيما إذا كان النموذج مناسباً او يحتاج الى تحسين وهل حقق النموذج الهدف المطلوب؟

2.3.3: التقلبات في الحمل الكهربائي

ذكر التقلبات العالم Mandelbrot في عام 1963 اذ أشار الى نتيجة مفادها "ان التغيرات الكبيرة تميل الى ان تتبعها تغيرات كبيرة"، وتميل التغيرات الصغيرة الى ان تتبعها تغيرات صغيرة". [55] كما يشكل التنبؤ بالتقلبات عملية بالغة الأهمية لقياس التقلبات في الحمل الكهربائي، فمحافظة

البصرة مرت بفترة زمنية من التقلبات، اذ شهد انتاج الطاقة الكهربائية تغيرات مختلفة. ففي المدة (1980-1990) شهدت الطاقة الكهربائية انخفاضا بفعل الحرب العراقية الإيرانية والعقوبات الاقتصادية المفروضة اذ تأثرت منظومة الطاقة الكهربائية في فترة الثمانينات بالحرب العراقية الإيرانية، اذ سجلت معدلات نمو سالبة في بداية المدة والتي بلغت (-0.20%) وبالرغم من تأثير الحرب والدمار الذي حل ببعض المحطات الكهربائية، الا انه تم انشاء العديد من المحطات لإنتاج الطاقة الكهربائية التي تميزت بانتشارها في المحافظات ويتفاوت توزيع هذه المحطات في المحافظات الشمالية والوسطى من العراق، ومن ثمَّ حققت معدلات النمو إيجابية في نهاية المدة 1989 والتي بلغت (1.31%). اما خلال المدة (1990 - 1999) فقد تأثرت منظومة الطاقة الكهربائية بأحداث حرب الخليج والحصار الاقتصادي المفروض على العراق والذي استمر حتى عام 2002، وخلال هذه المدة توقف تطور المحطات ولم تنشأ أي محطة توليد للطاقة الكهربائية في عموم العراق.

في عام 1991 تعرضت المحطات الكهربائية في البصرة وبقية مدن البلاد الى القصف الجوي والتدمير كما تأثرت المحطات بسبب العقوبات الاقتصادية المفروضة على العراق، فضلا عن ان معظم محطات الكهرباء قديمة ومستهلكة أدت الى نفاذ المعدات، وصعوبة توفير الأدوات الاحتياطية اللازمة لاستمرار تشغيل المحطات الكهربائية مما زاد من مشكلات القطاع الكهربائي وهذا إثر بشكل مباشر في الحياة الاقتصادية وحياء الافراد.

أدت الفوضى التي مر بها البلاد بعد 2003 والتدهور السياسي الى تدمير محطات توليد الطاقة الكهربائية وشبكتها، وتراجع المستوى التقني لمعدات وآلات المحطات الكهربائية. [23] بعد ذلك تم استيراد الطاقة الكهربائية من الدول المجاورة لسد العجز في الطلب وزيادة استهلاك الطاقة الكهربائية كما اعتمد المجتمع على المولدات الاهلية لتوفيرها وخاصة الطاقة الكهربائية المنزلية والتجارية. اذ وصل ما يستهلكه الفرد العراقي من الطاقة الكهربائية الى أدنى مستوياته وقد

إثر هذا الانخفاض في جميع الأنشطة العلمية والاجتماعية والزراعية والتجارية والصناعية لذا لابد من اتخاذ إجراءات خاصة لتأمين الطاقة الكهربائية للفرد. فتم انشاء عدد من المحطات الغازية فضلا عن توسعة بعض المحطات في جنوب العراق. [4] [24]

3.3.3: مفهوم التنبؤ (forecasting)

يعد التنبؤ من اهم أدوات اتخاذ القرار واهم عنصر في عملية الخطط المستقبلية، فمن اجل اتخاذ القرار الصحيح من الضروري دراسة جميع البدائل المتاحة وتحليل المتغيرات السابقة والحالية لتحديد الأفضل فالتنبؤ يعتمد على بيانات الماضي والحاضر من اجل معرفة المستقبل. [32]

للتنبؤ أهمية بالغة في التخطيط وصياغة القرارات الاقتصادية والاجتماعية وغيرها، اذ تبرز أهميته ودوره في الحد من المخاطر ويعطي صورة واضحة عن التوجه المستقبلي، فيضمن الكفاءة والفاعلية في المرونة مع البيئة الخارجية، ويسهم بقدر كبير في اتخاذ القرارات وترقب اثارها مستقبلا. فمن المهم استخدام الأساليب الإحصائية ذات الصلة للتخطيط ووضع الافتراضات حول الاحداث المستقبلية، بما في ذلك تقدير حجم الظاهرة في المستقبل مع الأخذ بنظر الاعتبار اهم العوامل التي تؤثر فيه. [19]

فالقرارات المتعلقة بالمستقبل تتخللها الكثير من المخاطر، ومن اجل تقليل هذه المخاطر نعتد على وظيفة تحليل الظواهر الماضية لتحديد الخصائص ومعرفة درجة التشابه بين الماضي والحاضر وهذه المعرفة ستؤدي الى زيادة الثقة في القرار المتخذ، ان مرحلة التنبؤ من اهم مراحل تحليل نماذج السلاسل الزمنية وهي الهدف الرئيس لعملية تقدير النموذج، اذ انه بعد تحديد النموذج وتقدير معلماته والتحقق منه وفحصه تأتي المرحلة الأهم وهي مرحلة التنبؤ اذ تتم معرفة سلوك الظاهرة قيد الدراسة في المستقبل. [33]

1.3.3.3 - خطوات بناء نموذج التنبؤ (Step of building forecasting model)

تدرج خطوات بناء نموذج مناسب للتنبؤ على وفق الآتي:

- تحديد النموذج (Model Identification) : يتم تمثيل السلسلة الزمنية ببيانيا، اذ ان الاحداثي الافقي يمثل الزمن والاحداثي الرأسي يمثل قيم الظاهرة (المشاهدات) ثم يتم اختيار نموذج رياضي بناء على بعض المقاييس الإحصائية لتمييز نموذج عن اخر وعلى الخبرة المستمدة من الدراسات والأبحاث.
- تقدير النموذج (Model Fitting) : بعد اختيار نموذج واحد او أكثر من النماذج المناسبة لوصف بيانات السلاسل الزمنية نقوم بتقدير معالم هذا النموذج من بيانات السلسلة الزمنية باستعمال طرائق التقدير الإحصائي.
- تشخيص واختبار النموذج (Model Diagnostics) : بعد تحديد النموذج والحصول على التقديرات يتم اجراء الاختبارات على البواقي لاختبار جودة النموذج ولمعرفة مدى صحة فرضيات النموذج ومدى تطابق المشاهدات مع القيم المحسوبة للنموذج المرشح، وفي حالة تحقق الشروط واختيار النموذج الافضل أي ان النموذج مناسب للبيانات يتم حساب التنبؤات بالقيم المستقبلية، اما إذا لم تتحقق فتتم إعادة الخطوات مرة أخرى حتى الوصول الى النموذج المناسب.
- ايجاد التنبؤات (Forecast Generation) : يتم استخدام النموذج النهائي لتوليد تنبؤات للقيم المستقبلية ثم حساب خطأ التنبؤ، عندما يتم ادخال قيم جديدة من السلسلة الزمنية ومراقبة هذه الأخطاء فيما يسمى بمخطط المراقبة اذ يتم وضع هذه الأخطاء للقبول بنسبة خطأ معين وإذا تجاوز خطأ التنبؤ يتم إعادة النظر في النموذج وتعاد الدورة من جديد لاختيار نموذج اخر.
- استخدام التنبؤات ووضع القرارات (Implementation and Decision making) : يتم تقديم التنبؤات لصناع القرار للنظر في استخدامها بالشكل المناسب. [7]، [3]

2.3.3.3 - أهمية التنبؤ (The importance of Forecasting) [26]

يعد التنبؤ اهم عنصر في عملية التخطيط ووضع الافتراضات المستقبلية باستخدام الأساليب الإحصائية التي ترسم مسار الظواهر المختلفة المراد التنبؤ بها، اذ تبرز أهمية ودور التنبؤ على وفق الآتي:

1. تسهم بقدر كبير في اتخاذ القرارات وتوقع تأثيرها في المستقبل.
2. يوفر للمنظمات صورة واضحة لاتجاهها المستقبلي.
3. يسهم في تقليل المخاطر التي قد تواجهها المنظمة.
4. معرفة حاجات المنظمة على المدى القصير والمتوسط.
- 5- يضمن والى حد كبير في فاعلية وكفاءه المنظمة في الحفاظ على المرونة مع البيئة الخارجية.

3.3.3.3 - خصائص التنبؤ (Forecast Process Properties) [32]

تتميز عملية التنبؤ بمجموعة من الخصائص نذكر منها الآتي:

1. التنبؤ عملية مستمرة ومرنة يتم تنفيذها تدريجياً.
 2. يعد التنبؤ طريقة شاملة تغطي جميع قضايا التنظيم.
 3. التنبؤ أسلوب مرحلي يتكون من مجموعة من الخطوات المتسلسلة والمتتابعة والمتراطة.
 4. يجب ان يكون التنبؤ طريقة متكاملة مع ابعاد الزمان والمكان ومحتواه وإجراءاته.
 5. التنبؤ هو اسلوب جماعي يسمح للمجموعة المستهدفة بالمشاركة.
- مفهوم التنبؤ رياضياً يمثل بافتراض ان (t) تشير الى المدة الزمنية الحالية التي يتم عندها حساب التنبؤات والمطلوب التنبؤ بقيمة المشاهدة التي ستحدث بعد (h) من المدد الزمنية، أي التنبؤ بقيمة المشاهدة (y_{t+h}) التي لم تحدث بعد. والحصول على اقل متوسط مربعات الخطأ للتنبؤ، تمثل (h) في هذه الحالة نطاق التنبؤ او افق التنبؤ (prediction – Horizon). وتشير $(y_t(h))$

الى القيمة التنبؤية التي نحصل عليها في المدة الزمنية (t) للمشاهدة (y_{t+h}) التي ستحدث بعد (h) من المدد الزمنية. فمثلا اذا كانت (h = 1)، اذن (y_t(1)) تشير الى القيمة التنبؤية التي تم الحصول عليها في المدة الزمنية (t) للمشاهدة (y_{t+1}) التي تحدث بعد مدة زمنية واحدة، وهكذا. بعد التأكد من ان النموذج المقترح يتصف بالدقة وملائم بشكل كاف، يتم استخدامه للتنبؤ بسلسلة المخرجات (y_t) في المستقبل بناءً على البيانات السابقة لكل من سلسلة المدخلات (x_t) وسلسلة المخرجات (y_t). [33]

4.3.3.3 - مقاييس دقة التنبؤ

توجد مجموعة متنوعة من المعايير التي يمكن استخدامها لاختبار دقة التنبؤ للنموذج المراد اختياره، فتعطي هذه المعايير الافضلية للنموذج الذي يعطي أدنى قيمة لها، ويتم ادراج هذه المعايير على وفق الاتي:

1. معيار متوسط مربعات الخطأ (Mean Square Error) : يمثل مجموع مربعات الخطأ مقسوما على عدد المشاهدات في السلسلة الزمنية وصيغته تكون على وفق الاتي:

$$MSE = \sum_{t=1}^n \frac{(y_t - \hat{y}_t)^2}{n} \dots (3.38)$$

حيث ان:

y_t: تمثل القيم الفعلية للمدة t ; \hat{y}_t : تمثل القيم التنبؤية للمدة t ; n: تمثل عدد المشاهدات في السلسلة الزمنية.

2. متوسط الخطأ المطلق (Mean Absolute Error): يمثل مجموع القيم المطلقة للخطأ

مقسومة على عدد المشاهدات في السلسلة الزمنية وتكون صيغته على وفق الاتي: [31]

$$MAE = \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{n} \dots (3.39)$$

3. متوسط الأخطاء النسبية المطلقة (Mean Absolute Percentage Error) : يمثل

مجموع حاصل قسمة القيمة المطلقة للخطأ مقسومه على القيمة الحقيقية مقسومه على عدد

المشاهدات في السلسلة الزمنية مضروبة في النسبة المئوية (100%) وتكون صيغته كالاتي: [31]

$$MAPE = \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} / n 100\% \dots (3.40)$$

حيث ان:

MAPE: هو مقياس الدقة في طراز السلسلة الزمنية ; N : تمثل عدد فترات التنبؤ.

4. الجذر التربيعي لمتوسط مربعات الخطأ (Root Mean Square Error): يقيس جودة التوافق بين البيانات الفعلية والنموذج المتوقع وتكون صيغته على وفق الاتي: [106]

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \dots (3.41)$$

وهناك معايير أخرى تستخدم للتحقق من دقة النموذج مقارنة مع النماذج المقترحة الأخرى ومنها:

1. معيار شوارتز Schwartz Bayesian Criterion : في عام 1978 اقترح الباحث

(Schwartz) معيارا مشابها لمعيار (Akaike) يسمى معيار (Schwartz) Bayesian

Criterion والذي يرمز له بالرمز (SBC) وصيغته على وفق الاتي:

$$SC(p) = \ln(\det|\sigma^2|) + \frac{2k^2 p \ln(N)}{N} \dots (3.42)$$

p : تمثل عدد الابطاء ; k : تمثل عدد المعلمات ; σ^2 : تمثل مصفوفة التباين والتباين المشترك

المقدرة للبقاي. [26]

ان دقة التنبؤات المتعلقة بالحمل الكهربائي في التخطيط المستقبلي قد حظيت باهتمام كبير في

النظم الاقليمية او الوطنية وخاصة في مجالات تخطيط تحميل الكهرباء والأنفاق على الطاقة

الكهربائية. الا ان التنبؤ بالحمل الكهربائي معقد وبيانات الاحمال الكهربائية تشكل اتجاهات غير

خطية للبيانات الناجمة عن العوامل المؤثرة. مثل العوامل الموسمية (تغير في المناخ الموسمي،

ونمو الاحمال) وعوامل المناخ (الرطوبة ودرجة الحرارة). إذ يؤدي التنبؤ بالحمل الكهربائي دورا

مهماً في توفير امدادات الطاقة الكهربائية الكافية دون انقطاع. [70]

الفصل الرابع

تقدير النموذج المختار والتنبؤ به

1.4: التحليل الوصفي للبيانات

2.4: المفهوم العام للظاهرة المدروسة

3.4: نتائج التحليل الاحصائي

3.4: المفاضلة بين النماذج المستخدمة في الدراسة

الفصل الرابع

تقدير النموذج والتنبؤ باستهلاك الطاقة الكهربائية

1.4: التحليل الوصفي للبيانات

ان البيانات التي تم اعتمادها في الدراسة الحالية هي بيانات سلسلة زمنية سنوية تمثل الاستهلاك السنوي للطاقة الكهربائية والمقدرة ب(ميكا واط، ساعة) (kwh) في المنطقة الجنوبية (البصرة، ميسان، ذي قار، المثنى) ولجميع أصناف الاستهلاك (التجاري، المنزلي، الحكومي، الصناعي، الزراعي، متجاوزين) التي تم الحصول عليها من الشركة العامة لتوزيع كهرباء الجنوب، اذ تمت دراسة المتغير التابع (Y) وتأثير اربعة متغيرات مستقلة (X_1, X_2, X_3, X_4) والتي تتكون كل سلسلة منها من (40) مشاهدة للسنوات من 1980 الى سنة 2019 .

1.1.4- وصف متغيرات الدراسة

الخطوة الأولى في اعداد النموذج تتمثل في تحديد المتغيرات المتضمنة فيه والتي يتم تحديدها من خلال الادبيات التي تناولت الظاهرة وعلى النحو الاتي:

المتغير التابع (Y): الاستهلاك السنوي للطاقة الكهربائية في المنطقة الجنوبية ويقاس ب(ميكا واط)

1- المتغيرات المستقلة:

X_1 : يمثل الناتج المحلي الإجمالي بالدينار العراقي (Gross domestic product).

X_2 : يمثل متوسط نصيب الفرد من الدخل بالدينار العراقي

(The average percapita income).

X_3 : يمثل عدد المشتركين في المنطقة الجنوبية (Number of participants).

X_4 : يمثل سعر الكيلو واط بالدينار/kwh (Kilowatt price in dinars).

2.1.4- الإحصاءات الوصفية

قبل التطرق الى التحليل الاحصائي يتم عرض بعض الإحصاءات الوصفية التي تمثل (الاستهلاك السنوي للطاقة الكهربائية، الناتج المحلي الإجمالي، متوسط نصيب الفرد من الدخل، عدد المشتركين والسعر) باستخدام برنامج (Matlab 7.11). لسلسلة زمنية تتكون من (40) مشاهدة وللمدة الزمنية (1980-2019) وتم الحصول عليها من مديرية توزيع كهرباء المنطقة الجنوبية ومركز السيطرة الجنوبي.

جدول (3) الإحصاءات الوصفية للمتغير المعتمد Y

Variable	N	Mean	Median	Min	Max
annual the consumption	40	6541883.63	3468755.50	2098227	24915560
	Range	Variance	Standard deviation	Kurtosis	Skewness
	22817333	4.410E13	6640909.415	1.978	1.837

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (spss)

يلاحظ في الجدول (3) ان الاستهلاك الكهربائي في المنطقة الجنوبية خلال مدة الدراسة يتراوح بين (24915560_2098227) بمتوسط حسابي (6541883.63) ومقدار التشتت عن وسطها الحسابي (6640909.415). كما يظهر في العمودين الأخيرين قيم معامل الالتواء والتفرطح لمتغير الاستهلاك السنوي للطاقة الكهربائية، اذ يتضح من قيم معامل الالتواء أنه ذو انحناء موجب اي ان التوزيع يكون ملتوياً التواءً موجباً نحو اليمين، اما قيمة التفرطح فنلاحظ من الجدول بانها قيمة موجبة أي ان المنحنى يكون منبسطة.

جدول (4) الإحصاءات الوصفية للمتغيرات التوضيحية

Variables	N	Mean	Median	Min	Max	Range
X_1	40	63524636.79	1878320	11012.6	268228000	268216987.4
X_2	40	1248226.245	81845.45997	34.295899	7645104.588	7645070.262
X_3	40	460784.25	405618.50	338081	825227	487146
X_4	40	24.0428950	2.7923000	0.0120	132.0588	132.0468
Variables	Variance	Standard deviation	Kurtosis	Skewness		
X_1	8.706E15	93306716.51	-0.570	1.045		
X_2	5.997E12	2448917.845	1.173	1.693		
X_3	1.670E10	129220.484	1.314	1.438		
X_4	1544.146	39.29561775	1.744	1.738		

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يمثل جدول (4) الإحصاءات الوصفية للمتغيرات التفسيرية للدراسة والتي تمثل (الناتج المحلي

الإجمالي (X_1)، متوسط نصيب الفرد من الدخل (X_2)، عدد المشتركين (X_3)، سعر الكيلو واط

(X_4)، نلاحظ في العمودين الأخيرين قيم معامل الالتواء (Skewness)، وقيم معامل التقطح

(Kurtosis)، إذ يظهر من قيم معامل الالتواء أن جميع المتغيرات تكون ذات انحناء موجب أي

أن منحنى التوزيع يكون ملتويًا التواءً موجباً نحو اليمين. أما معامل التقطح وكما موضح في

الجدول فنلاحظ أن قيم جميع المتغيرات موجبة ما عدا قيمة الناتج المحلي الإجمالي بالدينار

العراقي إذ تكون قيمته سالبة وهذا يعني أن شكل منحنى التوزيع منبسطة لجميع المتغيرات المستقلة.

وتجدر الإشارة إلى أن هذين المؤشرين يزودان الباحثة بأفكار أولية حول توزيع البيانات لتحديد ما

إذا كانت البيانات تتبع توزيعاً طبيعياً أو تحتاج البيانات إلى إجراء تعديلات وذلك لجعلها أقرب إلى

التوزيع الطبيعي.

2.4: المفهوم العام للظاهرة المدروسة

تعد الكهرباء أحد مصادر الطاقة في العصر الحديث وقد تعددت استعمالاتها حتى دخلت في جميع مجالات الحياة وأصبحت مقوماً من مقومات التقدم الاقتصادي، وتغطية الطلب المتزايد عليها في أي بلد تعد أحد أهم عناصر نجاح التنمية فيها. [28]

تعد الطاقة الكهربائية عنصراً حيوياً لتيسير الأعمال اليومية للمجتمع المعاصر. فتعد أحد أشكال الطاقة الرئيسية لبناء الحياة العصرية، لما لها من دور مهم في عملية الازدهار والتنمية الاقتصادية والثقافية نتيجة توسع استخدامها فهي متطلب مهم في الاقتصاد واستخدامها في تحسين وزيادة مستويات المعيشة للأفراد وتحسين دوره في تعزيز الأعمال المنزلية والخدمات الصحية والتعميم والسلامة وتقليل الأضرار البيئية مقارنة بأشكال الطاقة الأخرى. [2]

فالكهرباء عنصراً لا غنى عنه في حياة الإنسان، إذ تستخدم في الحياة المنزلية لتشغيل العديد من معدات التدفئة والتبريد والإضاءة وغيرها، فهي تستخدم لإقامة المشروعات الصناعية والتجارية والحكومية والزراعية. وقد أصبحت الطاقة الكهربائية من أهم مقومات التنمية في العديد من القطاعات الاستهلاكية مثل الصناعي والمنزلي والزراعي والحكومي والتجاري. [17]

كما عانت المناطق الجنوبية (البصرة، ذي قار، ميسان، المثنى) من أزمة في توفير الطاقة الكهربائية، فشهدت هذه المناطق انخفاضاً في الطاقة الكهربائية خلال المدة (2010 – 1990) ويعود هذا الانخفاض إلى توقف إنشاء المحطات لمدة (25) عاماً بسبب تأثير الحرب والحصار الاقتصادي، كذلك تعرض بعض المحطات الكهربائية في البصرة إلى التوقف المفاجئ نتيجة انخفاض مناسيب مياه شط العرب في عام 2010 وانخفاض كفاءه أجهزة تبريد الوحدات الكهربائية، فضلاً عن ذلك تعرض المحطات الكهربائية إلى أعمال السلب والنهب التي شهدها البلاد. [4]

هناك العديد من المشكلات في قطاع الكهرباء لا يمكن حلها الا من خلال الخطط الاستراتيجية في البلد. ومن اهم هذه المشكلات هي عدم وجود محطات توليد وخطوط نقل الطاقة خلال السنوات الأخيرة، فضلا عن ارتفاع معدلات الاحمال الكهربائية. كما ان انخفاض مستوى الطاقة الكهربائية يعود لأسباب عديدة، بما في ذلك فشل وزارة الكهرباء، وغياب التخطيط والمتابعة وعدم تجهيز الدوائر بالمعدات، وقلة محولات الإنتاج والتوزيع.[23]

يرتبط استهلاك الطاقة الكهربائية بدرجة وثيقة مع النمو الاقتصادي والاجتماعي ومدى تطور منظومة الطاقة الكهربائية من انتاج ونقل وتوزيع. اذ ارتفع استهلاك الطاقة الكهربائية في المنطقة الجنوبية في عام 2000 من (3504006) ميكاواط الى (6140592) ميكاواط خلال عام 2011. فيكون الاستهلاك متذبذبا لارتباطه بتذبذب الإنتاج اذ انخفض الاستهلاك خلال عام 2003 مقارنة مع عام 2002 بمقدار (201940) ميكاواط، كذلك في عام 2006 انخفض استهلاك الطاقة الكهربائية مقارنة مع عام 2005 بمقدار (8821) ميكاواط. اما فيما يتعلق باستهلاك الطاقة الكهربائية حسب القطاعات ففي عام 2000 تصدر القطاع الصناعي جميع القطاعات لتمتع المنطقة الجنوبية بأهم وأكبر الصناعات الأساسية وخاصة في محافظة البصرة مثل الصناعات النفطية والبتروكيماوية والاسمدة والحديد وصناعة الورق والمصافي وغيرها. وبعد انتهاء الحرب عام 2003 توقف العديد من القطاعات الصناعية بسبب المشكلات التي واجهها هذا القطاع مما أدى الى تراجع القطاع الصناعي الى المرتبة الثالثة بعد القطاعين الحكومي والمنزلي. في عام 2011 جاء القطاع المنزلي في المرتبة الأولى بنسبة 33.8% وبعده القطاع الحكومي بنسبة 31.3% اما القطاع التجاري والزراعي في المرتبة الأخيرة بنسبة 3.1% و0.9% على التوالي.[4]

3.4: نتائج التحليل الإحصائي

تم استخدام بيانات حقيقية لاستهلاك الطاقة الكهربائية في المنطقة الجنوبية للمدة من (1980) الى (2019) والتي تضمنت المتغير التابع (الاستهلاك السنوي للطاقة الكهربائية في المنطقة الجنوبية (ميكاواط، ساعة) والمتغيرات المستقلة التي تقيس تأثير نموذج انحدار المتجه الداعم هي كل من (الناتج المحلي الإجمالي بالدينار العراقي، متوسط نصيب الفرد من الدخل بالدينار العراقي، عدد المشتركين في المنطقة الجنوبية، سعر الكيلو واط بالدينار /kwh).

اقترح في هذا البحث طريقة مهجنة بين الخوارزميات الفوضوية (Chaotic algorithm) ونموذج انحدار المتجه الداعم (Support vector regression) لإيجاد أفضل المعلمات لنموذج انحدار المتجه الداعم.

وبعد اجراء عدة خطوات وبتكرارات متعددة تقوم الخوارزمية الفوضوية باختيار أفضل قيم للمعلمات (ϵ, δ, c) واختيار النموذج المناسب والأفضل للتنبؤ باستهلاك الحمل الكهربائي تم استخدام ثلاث

خوارزميات فوضوية وهي الخوارزمية الجينية الفوضوية (CGA) وخوارزمية المناعة الفوضوية

(CIA)، وخوارزمية تحسين سرب الطيور الفوضوية (CPSO) لتقدير معلمات انحدار المتجه

الداعم (SVR) وهي (C) في دالة الهدف (the objective function) وتمثل عامل تحديد

الخطأ إذ تحدد هذه المعلمة تكلفة المقايضة بين خطأ التدريب وتقليل تعقيد النموذج، كلما كانت

قيمة المعلمة (C) كبيرة فان الخطأ موجود، لذا يجب اختيارها بدقة لتجنب المبالغة في تقدير

التكاليف. ومعلمة (E) في دالة الخسارة (the loss function) والتي تستخدم لملاءمة بيانات

التدريب والتي تؤثر في عدد متجهات الدعم فكلما زاد حجم المعلمة يقل عدد متجهات الدعم.

ومعلمة (σ) في دالة النواه (the kernel function) اذ تستخدم هذه المعلمة لبناء دالة النواة

.Kernel

1.3.4- النموذج الأول: دمج انحدار المتجه الداعم (support vector regression) مع

الخوارزمية الجينية الفوضوية (chaotic genetic algorithm).

يتم تقدير معاملات انحدار المتجه الداعم (PSVR) مع الخوارزمية الجينية الفوضوية (CGA) وتحديد المعلمات (C, ϵ, σ) ، ثم استخدام أسلوب (cross – Valinda) لتقسيم البيانات بصورة مختلفة وقسمت الى ثلاثة فصول $k = 3$ (مجموعة التدريب ومجموعة التحقق من الصحة ومجموعة الاختبار) ليتم تقديرها. واستخدام الخوارزمية الجينية الفوضوية ومن خلال دالة اللياقة في الحصول على أفضل المعلمات. ويتم استخدام المعلمات الثلاثة (C, ϵ, σ) لنموذج (SVRCGA) التي تم تعديلها باستخدام (CGA) لحساب خطأ التحقق من الصحة، ثم يتم تحديد المعلمات التي تحتوي على الحد الأدنى من خطأ التحقق من الصحة كأكثر المعلمات ملائمة.

جدول (5) تقدير معلمات (SVR) مع الخوارزمية الجينية الفوضوية (CGA)

المعلمات الأمثل	متوسط مربع الخطأ	نموذج التقلب
δ	0.751	2.362
C	8.321×10^3	6.236×10^5
ϵ	0.66	0.82
عدد المدخلات	K=3	p=1, q=2
الخطأ العشوائي الأصغر	2.951	3.362

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يتضح من الجدول (5) تقدير معلمات انحدار المتجه الداعم (SVR) بطريقة الخوارزمية الجينية الفوضوية (CGA)، اذ تشير النتائج الى ان قيمة متوسط الخطأ لـ δ (0.751) وقيمته في النموذج (2.362)، وقيمة متوسط الخطأ لـ C (8.321×10^3)، وقيمته في النموذج (6.236×10^5)، كما ان قيمة متوسط الخطأ في ϵ (0.66)، وفي النموذج (0.82).

جدول (6) تقدير معاملات (SVR) مع الخوارزمية الجينية الفوضوية (CGA) بعد عملية التكرار

المعلمت الأمثل	متوسط مربع الخطأ	نموذج التقلب
δ	0.532	1.967
C	6.374×10^3	6.312×10^5
ε	0.56	0.78
عدد المدخلات	K=3	p=1, q=2
الخطأ العشوائي الأصغر	2.423	2.968

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يوضح الجدول (6) تقدير معالم نموذج انحدار المتجه الداعم (SVR) ومدى تأثير الخوارزمية

الجينية الفوضوية (CGA) بعد عملية تكرار اكثر من مره للحصول على تقنية افضل للنموذج

المهجن (CGASVR) من خلال اقل خطأ معياري، وقد أشارت النتائج الى ان قيمة متوسط الخطأ

لـ δ (0.532) وأيضاً في النموذج (1.967) وهذا يدل على ان الارتباط بين المتجه الداعم يكون

ضعيفاً، وان قيمة متوسط الخطأ لـ C (6.374×10^3) وفي النموذج (6.312×10^5) ويشير

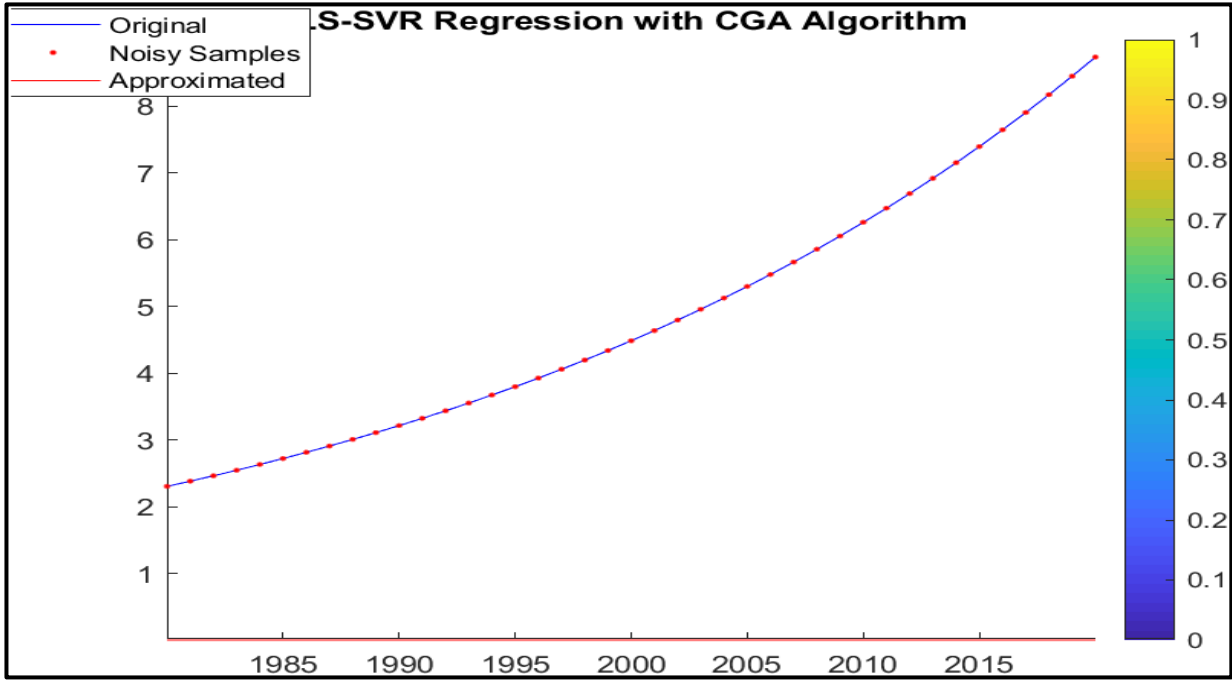
انخفاض قيمة C الى زيادة كفاءة انحدار المتجه الداعم (SVR) واستقرار نموذج التنبؤ. وان

متوسط الخطأ لـ ε (0.56) وفي النموذج (0.78)، والتي تؤدي الى زيادة تحسين دقة النموذج أي

ان نموذج التنبؤ معقد نسبياً. مما يؤكد لنا ان استخدام الخوارزمية الجينية الفوضوية (CGA) تؤثر

في نموذج انحدار المتجه الداعم (SVR) ولها تقنية عالية وفعالة من خلال قيم الخطأ المعياري

لأنه انخفض بعد تكرار العملية.



شكل (25) نتائج رسم انحدار المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA)

المصدر: من عمل الباحثة باستخدام برنامج (Matlab 7.11)

يوضح الشكل البياني (25) تأثير بناء نموذج انحدار المتجه الداعم (SVR)، اذ نجد ان هناك ارتفاعاً ملحوظاً في استهلاك الطاقة الكهربائية خلال المدة من (1980 – 1990) إذ يتوالى الارتفاع من عام (2010) وحتى نهاية المدة محل الدراسة (2019).

2.3.4- النموذج الثاني: دمج انحدار المتجه الداعم (support vector regression) مع

خوارزمية المناعة الفوضوية (chaotic immune algorithm).

يقترح هذا العمل ميزه جديدة حول خوارزمية المناعة الفوضوية (CIA) واهمية الانتروبيا في

المجموعة الأولية مستوحاة من نظرية الفوضى، تستخدم خوارزمية المناعة الفوضوية (CIA)

لتعديل المعلمات العشوائية، اذ تعدل الخوارزمية معلمات المجتمع الاولي باستخدام سلسلة فوضوية

ثم تحلل انتروبيا هذه المجموعة.

يتم تقدير معلمات انحدار المتجه الداعم (PSVR) مع خوارزمية المناعة الفوضوية (CIA)

وتحديدا المعلمات (C, ϵ, δ) ، اذ يتم استخدام أسلوب (cross – Valinda) لتقسيم البيانات

بصوره مختلفة ثم تقسم الى ثلاثة فصول $k = 3$ (مجموعة التدريب ومجموعة التحقق من الصحة ومجموعة الاختبار) ليتم تقديرها. واستخدام خوارزمية المناعة الفوضوية ومن خلال دالة اللياقة في الحصول على أفضل المعلمات. ويتم استخدام المعلمات الثلاثة (C, ϵ, σ) لنموذج (SVRCIA) التي تم تعديلها باستخدام (CIA) لحساب خطأ التحقق من الصحة، ثم يتم تحديد المعلمات التي تحتوي على الحد الأدنى من خطأ التحقق من الصحة كأكثر المعلمات ملائمة.

جدول (7) تقدير معلمات انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA)

المعلمات الأمثل	متوسط مربع الخطأ	نموذج التقلب
δ	0.831	2.634
C	9.435×10^3	6.785×10^5
ϵ	0.73	0.91
عدد المدخلات	K=3	p=1, q=2
الخطأ العشوائي الأصغر	1.96	3.362

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يتضح من الجدول المذكور أنفاً تقدير معلمات انحدار المتجه الداعم (SVR) باستخدام خوارزمية المناعة الفوضوية (CIA)، للحصول على تقنية افضل للتقدير من خلال اقل خطأ معياري، وتظهر النتائج ان متوسط خطأ لـ δ (0.831) وقيمة متوسط الخطأ لـ δ في النموذج (2.634)، وقيمة متوسط الخطأ لـ C تساوي (9.435×10^3) وقيمتها في النموذج (6.785×10^5) ، كما تشير الى ان متوسط الخطأ المعياري لـ ϵ (0.73) وفي النموذج (0.91).

جدول (8) تقدير معاملات انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA)

بعد عملية التكرار

المعاملات الأمثل	متوسط مربع الخطأ	نموذج التقلب
δ	1.963	3.251
C	4.263×10^3	6.952×10^5
ε	0.21	0.47
عدد المدخلات	K=3	p=1, q=2
الخطأ العشوائي الأصغر	1.89	3.842

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يوضح الجدول (8) تقدير معاملات نموذج انحدار المتجه الداعم (SVR) باستخدام خوارزمية

المناعة الفوضوية (CIA) بعد عملية تكرار اكثر من مره للحصول على تقنية افضل للنموذج

(CIASVR) من خلال اقل خطأ معياري، وقد اشارت النتائج الى ان متوسط الخطأ لـ δ

(1.963)، وكذلك قيمته في النموذج (3.251) ويدل ارتفاع قيمة المعلمة δ بعد عملية التكرار

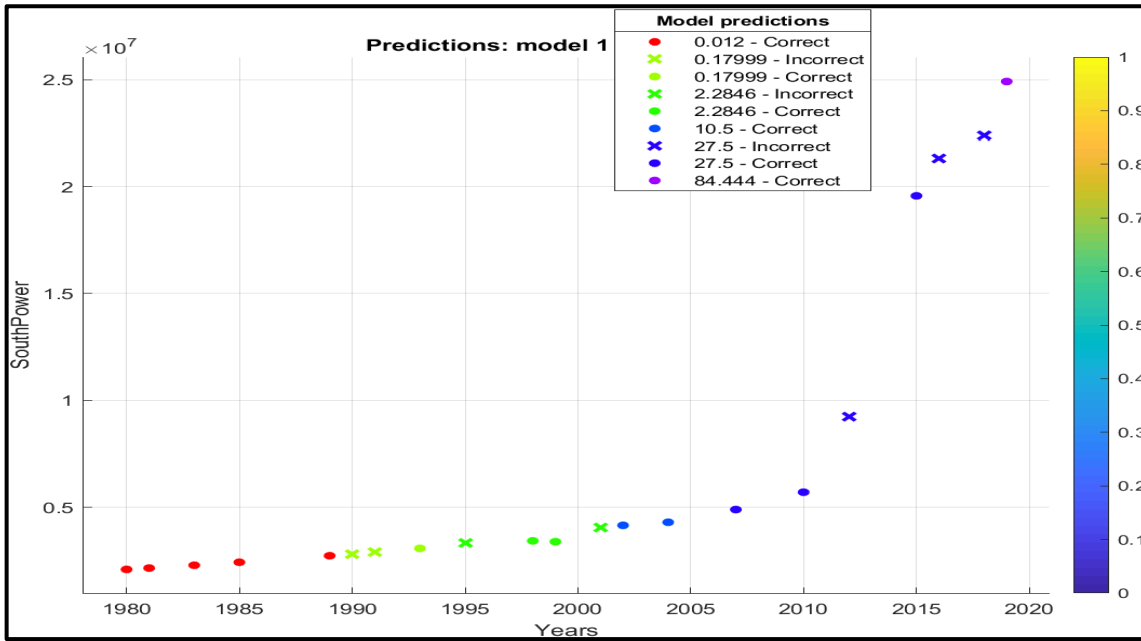
على ان الارتباط بين المتجه الداعم قوي جدا ولا يمكن الحصول على دقة كافية. وقيمة متوسط

الخطأ في C وفي النموذج (4.263×10^3)، (6.952×10^5) وتدل قيمة المعلمة C بعد عملية

التكرار الى ان أخطاء التدريب تكون كبيرة. ومتوسط الخطأ لـ ε (0.21) وأيضا في النموذج

(0.47) مما سيكون نموذج التنبؤ معقداً نسبياً، حيث ان قيمة الخطأ المعياري انخفضت بعد تكرار

العملية.



شكل (26) يوضح نتائج نموذج انحدار المتجه الداعم (SVR) مع خوارزمية المناعة الفوضوية (CIA)

المصدر: من عمل الباحثة باستخدام برنامج (Matlab 7.11)

يوضح الشكل البياني (26) تأثير بناء نموذج انحدار المتجه الداعم (SVR) من قبل خوارزمية المناعة الفوضوية (CIA)، إذ نلاحظ ارتفاع استهلاك الطاقة الكهربائية بعد عام 2010 وحتى نهاية المدة لبيانات الدراسة 2019.

3.3.4- النموذج الثالث: دمج انحدار المتجه الداعم (support vector regression) مع

خوارزمية تحسين سرب الطيور الفوضوية (Chaotic optimatizion the partial swarm)

يتم تقدير معاملات انحدار المتجه الداعم (PSVR) مع خوارزمية تحسين سرب الطيور الفوضوية

(CPSO) وتحديد المعلمات (C, ϵ, δ) ، إذ يتم استخدام أسلوب (cross - Valinda) لتقسيم

البيانات بصورة مختلفة ثم تقسيمها الى ثلاثة فصول $k = 3$ (مجموعة التدريب ومجموعة التحقق

من الصحة ومجموعة الاختبار) ليتم تقديرها. واستخدام خوارزمية تحسين سرب الطيور الفوضوية

ومن خلال دالة اللياقة في الحصول على أفضل المعلمات. ويتم استخدام المعلمات الثلاثة

(C, ϵ, σ) لنموذج (SVRCPSO) التي تم تعديلها باستخدام (CPSO) لحساب خطأ التحقق من

الصحة، ثم يتم تحديد المعلمات التي تحتوي على الحد الأدنى من خطأ التحقق من الصحة كأكثر المعلمات ملائمة.

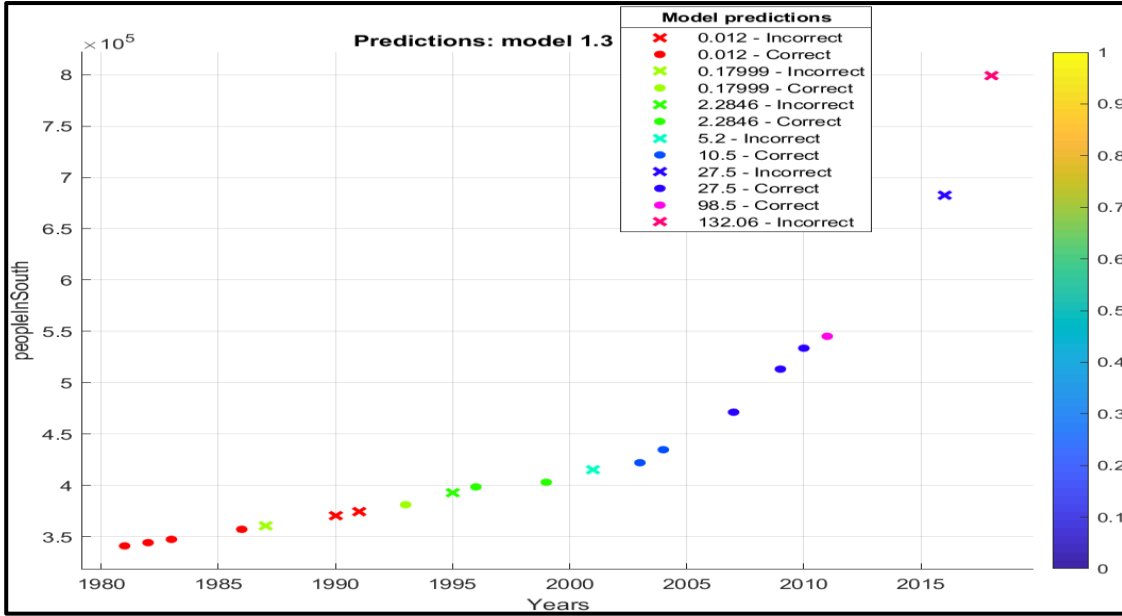
جدول (9) تقدير معلمات انحدار المتجه الداعم (SVR) مع خوارزمية تحسين سرب الطيور

الفوضوية (CPSO)

المعلمات الأمثل	متوسط مربع الخطأ	نموذج التقلب
δ	0.675	3.614
C	7.693×10^3	6.236×10^5
ϵ	0.74	0.67
عدد المدخلات	K=3	p=1, q=2
الخطأ العشوائي الأصغر	2.756	3.147

المصدر: من عمل الباحثة بالاعتماد على بيانات العينة باستخدام برنامج (Matlab 7.11)

يوضح الجدول المذكور أنفاً تقدير معالم نموذج انحدار المتجه الداعم (SVR) باستخدام خوارزمية تحسين سرب الطيور الفوضوية (CPSO)، بعد عملية تكرار أكثر من مره للحصول على تقنية أفضل لتقدير معلمات انحدار المتجه الداعم (SVR) من خلال اقل خطأ معياري. إذ اشارت النتائج الى ان متوسط الخطأ للمعلمة (δ) (0.675) وقيمتها في النموذج (3.614)، وهذا يدل على ان الارتباط بين متجهات الدعم ضعيفة ومن ثم تكون عملية التعلم الآلي معقدة نسبياً. وقيمة متوسط الخطأ لـ C (7.693×10^3) وفي النموذج (6.236×10^5)، وتشير هذه القيمة الى تحسين دقة التعلم. وان متوسط الخطأ المعياري لـ ϵ (0.74) وفي النموذج (0.67)، وتشير قيمة المعلمة (ϵ) الى ان المتجه الداعم يكون ضعيفا وهذا يؤدي الى ان نموذج التنبؤ بسيط وذو دقة منخفضة.



شكل (27) يوضح نتائج نموذج انحدار المتجه الداعم (SVR) مع خوارزمية تحسين سرب الطيور

الفوضوية (CPSO)

المصدر: من عمل الباحثة باستخدام برنامج (Matlab 7.11)

يتضح من النماذج السابقة ان كل نموذج مختلف عما سبقه بالنسبة لتحليل البيانات لذلك يتطلب النظر بدقة كافية في كل نموذج من النماذج الثلاثة حتى نتمكن من إيجاد افضل نموذج من حيث الدقة المطلوبة في تحليل البيانات وإيجاد المعلمات المرجو تحقيقها من الدراسة، لذلك وبالنظر الى النماذج المستخدمة في هذه الدراسة نجد ان لكل نموذج قيماً تدل على دقته وجودته في اختيار المعلمات المطلوبة لنموذج انحدار المتجه الداعم (SVR).

4.4: المفاضلة بين النماذج المستخدمة في الدراسة

يؤدي تطبيق انحدار المتجه الداعم (support vector regression) مع الخوارزمية التطورية (evolutionary algorithms) الى تحسين دقة التنبؤ بشكل كبير وتجنب التقارب السابق لأوانه بشكل فعال، فقد تمت المقارنة بين النماذج الثلاثة (SVRCGA،SVRCPSO،SVRCIA) على وفق معايير المفاضلة (MSE، MPE، MAPE، MAE) ومقياس (SCB) وكانت المعلمات التي تم الحصول عليها بواسطة النموذج الهجين لانحدار المتجه الداعم مع الخوارزمية

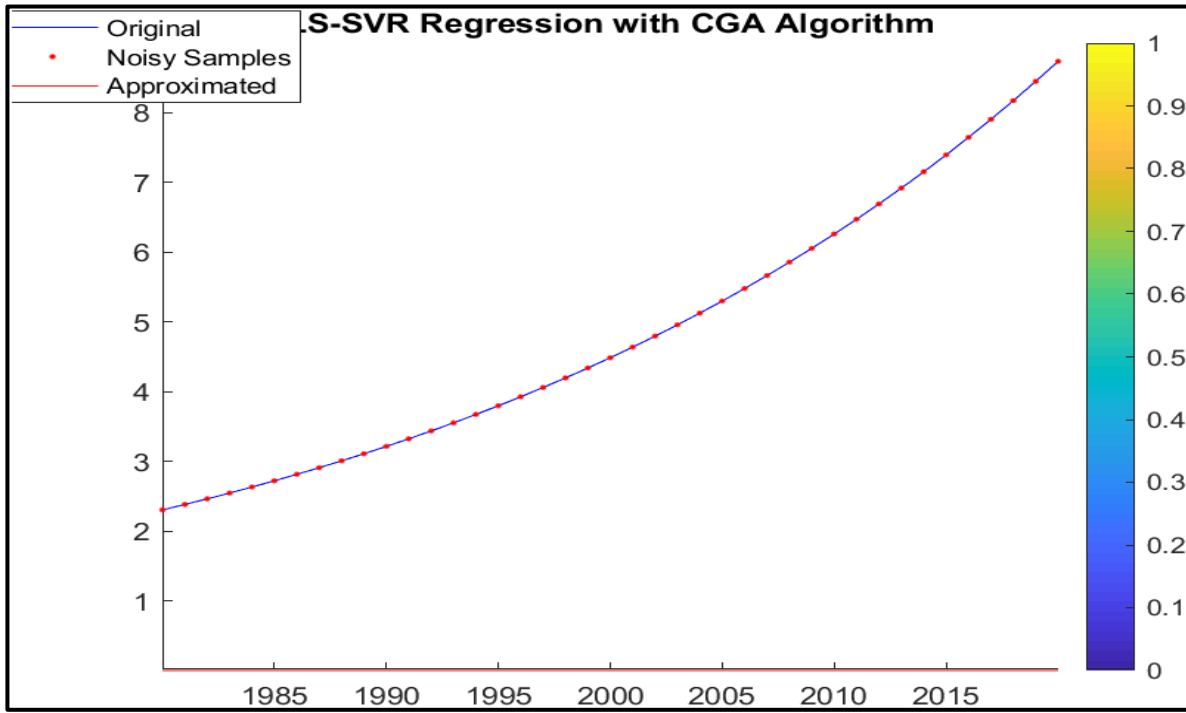
الجينية الفوضوية (SVRCGA) اكثر دقة من المعلمات المشتقة من نماذج (SVRCIA)،
(SVRCPSO). وكما موضح في جدول (15)

جدول (10) معايير المفاضلة بين النماذج

معايير لدقة التنبؤ	نموذج انحدار المتجه العام مع خوارزميه المناعة الفوضوية	نموذج انحدار المتجه العام مع خوارزميه سرب الطيور الفوضوية	نموذج انحدار المتجه العام مع خوارزميه الجينية الفوضوية
MSE	215389577888	6.007E11	1400000
RMSE	29460.0	245091.8195	1183.216
MAE	28830.0	350830.0	33.911
MAPE	7.415	1.000	0.12
SBC	4.6689	2.5339	2.0578

المصدر: من عمل الباحثة بالاعتماد على النتائج الاحصائية

يشير جدول (10) الى ان الخوارزمية الجينية الفوضوية (CGA) اكثر فعالية في تحسين معلمات انحدار المتجه الداعم (SVR)، فقد حققت دقة تنبؤ اكبر من الطرائق الأخرى، لذا يتم اختيار النموذج الهجين لانحدار المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA) الذي تم الحصول عليه لفعاليتها العالية في التنبؤ واستخدامه للتنبؤ ببيانات الاختبار في استهلاك الحمل الكهربائي في المنطقة الجنوبية من العراق .



شكل (28) نتائج المفاضلة بين النماذج الثلاثة

المصدر: من عمل الباحثة باستخدام برنامج (Matlab 7.11)

يشير الشكل (28) الى ان نموذج (SVRCGA) المقترح يحقق نتائج تكون أقرب الى قيم الحمل

الكهربائي الفعلية من النماذج الأخرى، ويشير أيضا إلى أن نموذج (SVRCGA) المقترح يوفر

مساهمات كبيرة في التحسينات من حيث دقة التنبؤ باستهلاك الطاقة الكهربائية في المنطقة

الجنوبية.

5.4: نتائج التنبؤ (Forecasting Results)

باستخدام النموذج الذي تم الحصول عليه (SVRGA)، فقد تم التنبؤ بكميات الاستهلاك السنوي

للطاقة الكهربائية في المنطقة الجنوبية للمدة من 2020 الى 2028 أظهرت النتائج ان النموذج

(CGASVR) يتفوق على النماذج الأخرى إذ يمكن ان نلاحظ من الجدول (10) ان نموذج انحدار

المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA) يمتلك اصغر قيم

(MSE، MAPE، MAE، RMSE) بين النماذج الأخرى في هذه الدراسة. وكذلك يبين الجدول

(11) التنبؤات بالحمل الكهربائي للمدة من (2020 – 2028) في المنطقة الجنوبية (البصرة، ذي قار، ميسان، المثنى). وتم رسم السلسلة الزمنية لهذه التنبؤات وكما يظهر في الشكل (4).

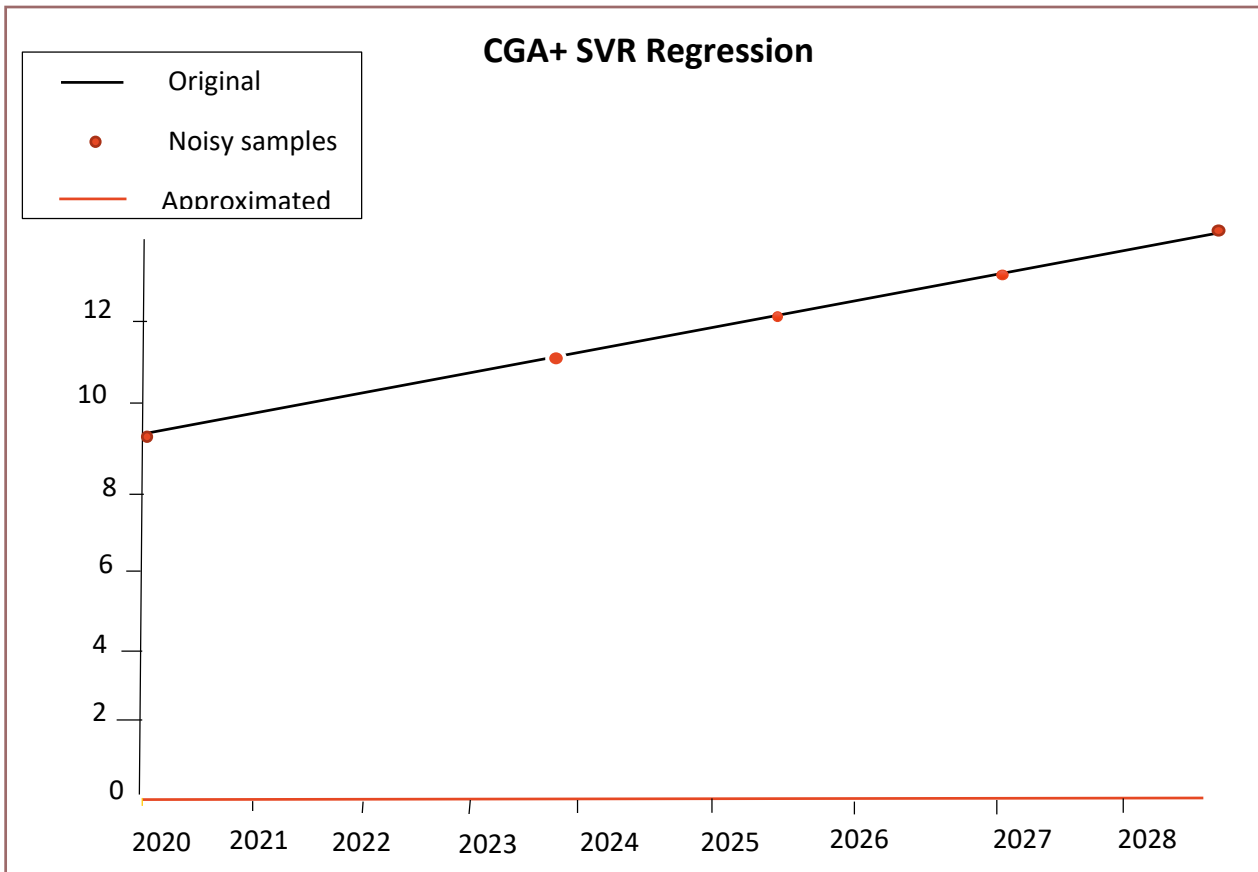
جدول (11) قيم التنبؤ المستقبلية باستهلاك الطاقة الكهربائية (2020 – 2028)

السنوات	قيم التنبؤ باستهلاك الحمل الكهربائي باستخدام خوارزمية (CGA+SVR) (جيجا واط/ساعة)
2020	9236057
2021	12790829
2022	16353621
2023	19568089
2024	21317833
2025	22847328
2026	22389174
2027	24915560
2028	9236057

المصدر: من عمل الباحثة اعتماداً على بيانات العينة وباستخدام برنامج (Matlab 7.11)

يمكن رسم السلسلة الزمنية المتنبأ بها لـ 9 سنوات أي من سنة 2020 إلى سنة 2028، ويوضح

الشكل (29) رسم القيم التنبؤية لسلسلة بيانات الاستهلاك السنوي للطاقة الكهربائية.



شكل (29) رسم القيم التنبؤية لسلسلة بيانات الاستهلاك السنوي للطاقة الكهربائية (2028-2020)

المصدر: من عمل الباحثة باستخدام برنامج (Matlab 7.11)

نلاحظ من الشكل (29) إن الاستهلاك سيتم في ازدياد لكن بنسب ثابتة على وفق العوامل

المستخدمة في الدراسة وبناء على استخدام النموذج المقترح (SVRCGA)، إذا استخدمت بالشكل

المناسب على وفق خطط مدروسة.

المصادر

أولاً: المصادر العربية

1. أبو زيد، أماني محمد (2019). "فاعلية وحدة معدلة وفق النمذجة المفاهيمية في تنمية مهارات الجدل العلمي في تعلم البيولوجي والتفكير التنسيقي لدى طالب المرحلة الثانوية"، كلية التربية، جامعة عين الشمس، مجلة كلية التربية، العدد الثالث والاربعون (الجزء الثالث).
2. أبو صبيح، مازن سلطان عزيز (2007). "إنتاج الطاقة الكهربائية لتلبية الطلب الكلي في العراق: دراسة تحليلية للمدة (1970- 1999) "، أطروحة دكتوراه، جامعة البصرة، العراق.
3. أبو لبة، محمد عبد الرحمن جاد الله (2017). "استخدام نماذج ARMAX في التنبؤ بالسلاسل الزمنية"، كلية الاقتصاد والعلوم الإدارية، جامعة الأزهر، غزة. الأساسية المجلد (36) العدد (5).
4. الأسدي، كاظم عبد الوهاب حسن والشريفي، راشد عبد راشد (2015). "صناعة الطاقة الكهربائية في جنوب العراق"، مجلة الخليج العربي، مج. 43، ع. 3-4، ص50-95.
5. بحوح، اسامه سعد وريشة، حسان (2007). "تأثير متغيرات الخوارزمية الجينية في مسائل إيجاد الحل الأمثل"، مجلة جامعة دمشق للعلوم الهندسية، المجلد الثالث والعشرون، العدد الثاني، ص 105-128.
6. البدراني، مها عبد الاله محمد (2007). "استخدام الخوارزمية الجينية في تطابق أنماط الحرف الإنكليزي"، مجلة التربية والعلم - المجلد (19) العدد (4).
7. بري، عدنان ماجد (2002)، "طرائق التنبؤ الاحصائي". الجزء الأول، كتاب، جامعة الملك سعود، رياض.
8. بشير، عالية توفيق (2018). "تقدير معلمات نموذج الانحدار المبتور باستخدام خوارزميتي سرب الطيور وأشباه نيوتن"، المجلة العراقية للعلوم الإحصائية، مج. 15، ع. 27، ص. 83-106.
9. بلال، وسيم حبيب (2014). "استخدام الطرق الهجينة للمساهمة في حل مسألة توجيه المركبة"، جامعة تشرين، مجلة جامعة تشرين للبحوث والدراسات العلمية - سلسلة العلوم
10. توبة، رباب احمد عبد القادر (2014). "إثر استخدام استراتيجيات النمذجة الرياضية على استيعاب المفاهيم الرياضية وحل المسألة الرياضية لدى طلبة الصف السابع الأساسي في وحدة القياس"، كلية الدراسات العليا، جامعة النجاح الوطنية.
11. ثابت، همسه معن محمد (2005). "بعض تطبيقات الخوارزمية الجينية في حل مسائل الامثلية"، جامعة الموصل، كلية علوم الحاسبات والرياضيات.
12. الثعلبي، ساهرة حسين زين والثعلبي، سكنه جهيه فرج وخزعل، ريسان عبد الامام. (2018). "الاستثمارات العربية البنينة ودورها في تحقيق النمو الاقتصادي للمدة (1981-2017) - دراسة قياسية" جامعة البصرة، العراق، المجلة العالمية للاقتصاد والأعمال (5)، العدد (3). ص342-361.

13. الحسني، فاتن محمد علي (2014). "إثر استخدام النمذجة الرياضية على تنمية مهارات التفكير المنظومي في الرياضيات والميل نحوها لدى طالبات الصف الخامس الأساسي بغزة"، قسم المناهج وطرائق التدريس، كلية التربية، الجامعة الإسلامية، غزة.
14. حمادي، صباح سعيد (2016). "إثر أسلوب النمذجة الرياضية في حل المشكلات التطبيقية لدى طلاب الصف الثاني متوسط في الرياضيات"، مجلة الأستاذ، العدد 217، المجلد الثاني، (305-338).
15. الخالدي، خليل إبراهيم عيسى (2017). "دور خدمات الامداد الإنساني في تحسين فاعلية ادارة الكوارث: مرتكزات نظرية الفوضى متغير تفاعلي"، قسم إدارة اعمال، كلية الإدارة والاقتصاد، جامعة البصرة.
16. الخفاجي، سحر حسين (2016). "فاعلية استراتيجية النمذجة في تحصيل طالبات الصف الرابع الادبي في مادة التاريخ". مجلة الفتح، العدد السابع والستون.
17. الشريفي، راشد عبد راشد (2013). "التوزيع الجغرافي لإنتاج واستهلاك الطاقة الكهربائية في العراق"، أطروحة دكتوراه، قسم الجغرافية، كلية الآداب، جامعة البصرة.
18. الصمد، بوشنه عبد. (2016). " اختبار علاقة التكامل المشترك لأثر التغير في التداول النقدي على الناتج الداخلي الخام دراسة حالة الجزائر خلال الفترة 1970-2014". جامعة قاصدي مرباح ورقلة، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، قسم العلوم الاقتصادية.
19. الطائي، فاضل عباس (2010). "التنبؤ والتمهيد للسلاسل الزمنية باستعمال التحويلات مع التطبيق"، بحث منشور في مجلة العراقية للعلوم الاقتصادية، ص 293-308، العدد 17.
20. الطفيلي، مهدي عبد الأمير والعوادي، ضرغام سعد فليح (2020). "فاعلية استخدام طريقة النمذجة في إتقان النسب التشريحية للبورترية في مادة التخطيط"، مجلة مركز بابل للدراسات الإنسانية مج. 10، العدد 1، ص 187-208.
21. عبد الجليل، إبراهيم وشاكر، أمجد (2018). "دراسة تأثير تغيير بارامترات الخوارزمية الجينية الفوضوية على الحلول الأمثلية في تصميم شبكات مياه الشرب"، مجلة جامعه حماة، المجلد الأول، العدد الثاني.
22. عمر، موفق (2019). "استخدام الخوارزميات الجينية في التنبؤ بتطايير الأسواق المالية"، قسم التقنيات الكمية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة أبي بكر بلقايد - تلمسان.
23. عمران، خلود موسى. الثعلبي وساهرة، زين (2015). "تقدير الطلب على الكهرباء في محافظة البصرة للمدة 1995-2010"، مجلة الغري للعلوم الاقتصادية والإدارية، المجلد التاسع، العدد 32، ص (104-133).
24. عودة، محمد حسن واحمد، حسين علي (2016). واقع إنتاج الطاقة الكهربائية في مدينة البصرة: نشوء، التطور، المعوقات"، مجلة الغري للعلوم الاقتصادية والإدارية، مج. 13، العدد 39، ص. (47-73).

25. عيسى، خليل إبراهيم (2017). "تأثير الخدمات اللوجستية الإنسانية في تحسين فاعلية ادارة الكوارث في إطار الدور التفاعلي لمركزات نظريه الفوضى: دراسة تطبيقية في وزاره الهجرة والمهجرين والمنظمات الدولية الانسانية في العراق"، مجلة الاقتصاد الخليجي، العدد (32)، ص (17-22).
26. الكلابي، صفاء مجيد مطشر (2018). "استعمال بعض طرائق التنبؤ المختلفة لتحليل اعداد المصابين بالأورام الخبيثة"، قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة كربلاء.
27. الكلاك، اسراء نذير وشعبان، رقية زيدان (2008). "الخوارزمية الجينية في جدولة العمليات مع عدم إمكانية القطع"، مجلة تنمية الرافدين، العدد 30، ص (245-256).
28. المجالي، احمد عبد القادر والرفوع، احمد سلمان (2018). "تقدير الطلب على استهلاك الطاقة الكهربائية للقطاع المنزلي في الأردن باستخدام نموذج تصحيح الخطأ (VECM) للفترة (1980 – 2015)"، المجلة الأردنية للعلوم الاقتصادية، المجلد (5)، العدد (1).
29. مجموعة من الباحثين (2019). "تطبيقات الذكاء الاصطناعي كتوجه حديث لتعزيز تنافسية منظمات الأعمال". المركز الديمقراطي العربي للدراسات الاستراتيجية والسياسية والاقتصادية، برلين، ألمانيا، الطبعة الأولى.
30. محمد، صالحة رائد وصالح، إبراهيم احمد (2018). "استخدام خوارزمية أمثلية عناصر السرب لقياس جودة البرمجيات"، مجلة الرافدين لعلوم الحاسبات والرياضيات، المجلد (12)، العدد (1)، ص (26-36).
31. محمود، غزوان هاني (2010). "تحسين طريقة التمهيد الاسي البسيط للتكهن بالسلاسل الزمنية"، المجلة العراقية للعلوم الإحصائية (18) ص 259-272.
32. مراس، محمد وحديد، زهية (2017). اهمية نمذجة التنبؤ باستخدام نماذج السلاسل الزمنية على معطيات اتخاذ القرار في المؤسسة الاقتصادية الجزائرية _ دراسة حالة القطاع الخدماتي"، جامعة المسيلة، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التيسير.
33. مطر، ظافر رمضان والياس، انتصار إبراهيم (2010). "تحليل ونمذجة السلسلة الزمنية لتدفق المياه الداخلة الى مدينة الموصل دراسة مقارنة"، المجلة العراقية للعلوم الإحصائية، العدد 18، ص (32-1).
34. الوزان، محمد علي محمد وقاسم، عمر صابر (2018). "تحسين أداء تصنيف تقانة آلة المتجه الداعم باستخدام الخوارزمية الجينية"، كلية علوم الحاسوب والرياضيات جامعة الموصل، العراق، مجلة الرافدين لعلوم الحاسوب والرياضيات، المجلد (12) العدد (2).
- ثانياً: المصادر الإنكليزية

35. Al. Anie, R.M., Al. Quraishy, H.A. (2011). "Partial swarm optimization method", University of Baghdad, College of Science, Department of Computer science, pp (1-27).

36. Alhafedh, M.A.A., Qasim, O.S. (2019). "A Proposed Method for Feature Selection using a Binary Particle Swarm Optimization Algorithm and Mutual Information Technique", College of Computer Science and Mathematics University of Mosul, Iraq, Volume (13), No. (2), pp (51-52).
37. Al-Sa'edi, M., Alwan, H.A. (2014). "Organizational dialectical convergence chaos and complexity theories in Iraqi business organizations (experimental study in a sample of community colleges)", Journal of Techniques, Volume 27, Issue 5, pp (23-43).
38. Atilgan, E. (2016). "Computational Doping for Fuel Cell Material Design Based on Genetic Algorithms and Genetic Programming", University of South Carolina, (Doctoral dissertation).
39. Bai, Q. (2010). "Analysis of Particle Swarm Optimization Algorithm", College of Computer Science and Technology Inner Mongolia University for Nationalities, China, Computer and Information Science, Vol .3, No. 1.
40. Bajpal, p. (2008). "Genetic Algorithm – an Approach to Solve Global Optimization Problems", Amity Institute of Information Technology, Amity University, Lucknow, Indian Journal of Computer Science and Engineering Vol .1 ,No .3 ,pp (199-206).
41. Balasubramaniam, P. (2011). "Control, Computation and Information Systems" , Department of Mathematics ,Gandhigram Rural Institute ,India, pp (1-328), Springer Heidelberg Dordrecht London New York.
42. Bonfil, G.S., Solís, J.F. & Rodarte, I.V. (2013). "Volatility Forecasting Using Support Vector Regression and a Hybrid Genetic Algorithm", Computational Economics, 45(1), pp (111–133).
43. Cai, J., Ma, X., Li, L. & Haipeng, P. (2007). "Chaotic particle swarm optimization for economic dispatch considering the generator constraints", Energy Conversion and Management, 48(2) , pp (645–653).
44. Carmen, P. L., & Ricardo, L. R. (2012). "Notions of Chaotic Cryptography: Sketch of a Chaos Based Cryptosystem", Department of Computer Science and

- BIFI, University of Zaragoza, Spain.
45. Chen, C. W. & Chang, Y. C. (2016). "Support Vector Regression and Genetic Algorithm for HVAC Optimal Operation", National Taipei University of Technology, Hindawi Publishing Corporation, Volume 2016, No. 1
46. Chen, K. Y., & Wang, C. H. (2007). "Support vector regression with genetic algorithms in forecasting tourism demand", *Tourism Management*, 28(1), pp (215–226).
47. Chen, P., Cuzzocrea, A., Du, X., Kara, O., Liu, T., Ślęzak, D. and Yang, X. (2018). "Advanced Informatics for Computing Research ", Second International Conference, India, Revised Selected Papers, Part I.
48. Chen, L. Y.; Hong, W. C.; Panigrahi, b. K.; and Wei, S. Y. (2011). SVR with Chaotic Genetic Algorithm in Taiwanese 3G Phone Demand Forecasting, Part I, pp. (248–256), Springer-Verlag Berlin Heidelberg.
49. Cheng, Y., Fu, Y., & Gong, G. (2010). "A Hybrid Support Vector Regression Based on Chaotic Particle Swarm Optimization Algorithm in Forecasting Financial Returns", School of Computer Science & Technology, Soochow University, China.
50. Coelho, L. d. S. (2008). "A quantum particle swarm optimizer with chaotic mutation operator", *Chaos, Solitons & Fractals*, vol. 2008, No.37, pp (1409–1418).
51. Cui, H., & Song, X. (2008). "An Empirical Research on Short Term Power Load Forecasting Based on Chaos Theory", Department of Economics and Management, North China Electric Power University, China, pp (394-397).
52. Deep, K., Nagar, A., Pant, M. & Bansal, J.C. (2012). "Advances in Intelligent and Soft Computing", Proceedings of the International Conference on Soft Computing for Problem Solving", Volume 1, pp (933-941), Springer New Delhi Heidelberg New York Dordrecht London.
53. Deng, N., Tian, Y., Zhang, C. (2013). "Support Vector Machines Optimization Based Theory Algorithms and Extensions", University of Minnesota, Department

- of Computer Science and Engineering Minneapolis, U.S.A, pp (1-315).
54. Deng, Y., Xiang, J. & Ou, Z. (2012). "SVR with hybrid chaotic genetic algorithm for short-term traffic flow forecasting", Mathematical department, School of science, Wuhan University of Technology, pp (708-712).
55. Ding, D. (2011). "Modeling of Market Volatility with APARCH Model", Department of Mathematic, Uppsala University, pp (1-42).
56. Dong, Y., Zhang, Z. & Hong, W. C. (2018). "A Hybrid Seasonal Mechanism with a Chaotic Cuckoo Search Algorithm with a Support Vector Regression Model for Electric Load Forecasting", *Energies*, vol.11, No. 4, pp (2-21).
57. Duan, K., Keerthi, S.S. & Poo, A. N. (2003). "Evaluation of simple performance measures for tuning SVM hyperparameters", Department of Mechanical Engineering, National University of Singapore, pp (41-59).
58. Fan, S., & Chen, L. (2006). "Short-Term Load Forecasting Based on an Adaptive Hybrid Method", *IEEE Transactions on Power Systems*, vol.21, No. 1, pp (392–401).
59. Getchell, M. (2018). "Chaos Theory and Emergent Behavior in the West Virginia Water Crisis", *journal of international crisis and risk* ,vol. 1, No. 2, pp (173–200).
60. Gunn, S. R. (1998). "Support vector machines for classification and regression", *journal The Royal Society of Chemistry*, Vol. 2010, No.135, pp (230-267)
61. Hammad, M. A., Jereb, B., Rosi, B. & Dragan, D. (2020). "Methods and Models for Electric Load Forecasting: A Comprehensive Review", *Logistics & Sustainable Transport*, vol.11, No. 1, pp (51–76).
62. Haniyas, M.P., Curtis, P.G. and Thalassinos, Y.E. (2009). "Chaos theory: forecasting the freight rate of an oil tanker", *Int. J. Computational Economics and Econometrics*, Vol. 1, No. 1, pp (76-85).
63. Hnin, Su wutyi. (2017). "Hybrid support vector regression for short term electricity load forecasting", master of science engineering and technology , sirindhorn international institute of technology, thammasat university.

64. Hong, W. C. (2008). "Rainfall forecasting by technological machine learning models", *Applied Mathematics and Computation*, vol. 200, No. 1, pp (41–57).
65. Hong, W. C. (2009). "Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model", *Energy Conversion and Management*, vol.50, No. 1, pp (105–117).
66. Hong, W. C. (2009). "Electric load forecasting by support vector model", *Applied Mathematical Modelling*, vol. 33, No. 5 , pp (2444–2454).
67. Hong, W. C., Dong, Y., Chen, L. Y. & Wei, S. Y. (2012). "Seasonal Support Vector Regression with Chaotic Genetic Algorithm in Electric Load Forecasting", *Sixth International Conference on Genetic and Evolutionary Computing*, pp (124-127)
68. Hong, W. C., Dong, Y., Chen, L. Y., & Wei, S. Y. (2011). "SVR with hybrid chaotic genetic algorithms for tourism demand forecasting", *Applied Soft Computing*, vol. 11, No. 2 , pp (1881–1890).
69. Hong, W. C., Dong, Y., Lai, C. Y., Chen, L. Y. & Wei, S. Y. (2011). "SVR with Hybrid Chaotic Immune Algorithm for Seasonal Load Demand Forecasting ", *Energies*, vol. 4, No. 6, pp (960–977).
70. Hong, W. C., Dong, Y., Zhang, W. Y., Chen, L. Y. & Panigrahi, B. k. (2013). "Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm", *International Journal of Electrical Power & Energy Systems*, vol. 44, No. 1, pp (604–614).
71. Hong, W. C., Lee, S. L., Lai, C. Y., Wu, Y. H. & Wang, K. L. (2007). "The potentiality of support vector regression with immune algorithm for regional electric load forecasting", *International Joint Conference on Neural Networks*, No.58, pp (12-17).
72. Hong, W.C. (2013). "Intelligent Energy Demand Forecasting", Department of Information Management ,Oriental Institute of Technology , City Taiwan.

73. Ionescu, L. M., Bizon, N., Mazare, A. G., & Belu, N. (2020). "Reducing the Cost of Electricity by Optimizing Real-Time Consumer Planning Using a New Genetic Algorithm-Based Strategy", *Mathematics*, vol.8, No. (7), pp (2-26).
74. Jaafri, Y.O. (2016). "Recognition of Printed Arabic Script Using Support Vector Machine", University Ahmed Draia - Adrar, Faculty of Sciences and Technology, Department of Mathematics and Computer Science (Master thesis) .
75. Jakkula, V. (2003). "Tutorial on Support Vector Machine (SVM)", School of EECS, Washington State University.
76. Jiang, H., Yu, X. & Yu, Y. (2008). "Application of Support Vector Regression in Power System Short Term Load Forecasting", *Fourth International Conference on Natural Computation*, pp (26-30).
77. Jingjing, W., Zheng, Y., Qi, W. & Shan, F. (2014). "Application of Chaotic Genetic Algorithm and Elman to Assess Flight Performance Under Multiple Physiological Signals", *The Open Automation and Control Systems Journal*, vol. 6, pp (975-983).
78. Ju, F. Y. & Hong, W. C. (2013). "Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting. *Applied Mathematical Modelling*", vol. 37, No. 23, pp (9643–9651).
79. Kavaklioglu, K. (2011). "Modeling and prediction of Turkey's electricity consumption using Support Vector Regression", *Applied Energy*, vol. 88, No. 1, pp (368–375).
80. Kavzoglu, T., & Colkesen, I. (2009). "A kernel functions analysis for support vector machines for land cover classification". *International Journal of Applied Earth Observation and Geoinformation*, vol.11, No. (5), pp (352–359).
81. Kawauchi, S., Sugihara, H. & Sasaki, H. (2004). "Development of very-short-term load forecasting based on chaos theory", *Electrical Engineering in Japan*, vol. 148, No. 2 , pp (55–63).
82. Lahiri, S. K., & Khalfe, N. (2009). "Process modeling and optimization of industrial ethylene oxide reactor by integrating support vector regression and

- genetic algorithm", The Canadian Journal of Chemical Engineering, vol. 87, No. 1, pp (118–128).
- 83.Lahiri. S. K., Ghanta, K. C. (2008)., "The Support Vector Regression with the parameter tuning assisted by a differential evolution technique: Study of the critical velocity of a slurry flow in a pipeline", Chemical Industry & Chemical Engineering Quarterly, vol. 14, No. 3, pp (191–203).
- 84.Langhammer, J., Cesák, J. (2016). "Applicability of a Nu-Support Vector Regression Model for the Completion of Missing Data in Hydrological Time Series, Water, vol. 8, No. 560, pp (2-25).
- 85.Laura, A., Rouslan, M. (2008). "Support Vector Machines (SVM) as a technique for solvency analysis, SSRN Electronic Journal, No. 811, pp (1-16).
- 86.Lee, C.W., Lin, B.Y. (2017). "Applications of the Chaotic Quantum Genetic Algorithm with Support Vector Regression in Load Forecasting", *Energies*, vol.10, No. 11, pp (2-18).
- 87.Lee, K.H. (2005). " First Course on Fuzzy Theory and Applications", Advanced Institute of Science and Technology, Republic of South Korea, Springer Berlin Heidelberg New York.
- 88.Li, M. W., Geng, J., Wang, S. & Hong, W. C. (2017). " Hybrid Chaotic Quantum Bat Algorithm with SVR in Electric Load Forecasting ", *Energies*, vol.10, No. 12, pp (2-18).
- 89.Li, T., Yong.W., Zhi, GZ., Zhi, L., Chen, YG. (2014). "Short-Term Wind Power Prediction Based on Support Vector Regress Optimized by Chaos Particle Swarm Optimization", School of Electrical Engineering, Anhui Polytechnic University, China, pp (195-200).
- 90.Li, W., Sheng, D., Chen, J., Yuan, Z.& Cen, K. (2006). "anew method based on immune algorithm to solve the unit commitment problem ", College of Mechanical and Energy Engineering, Zhejiang University, Vol.207, pp (840-846).

91. Lin, K. P., Pai, P. F. & Yang, S. L. (2011). "Forecasting concentrations of air pollutants by logarithm support vector regression with immune algorithms", *Applied Mathematics and Computation*, vol. 217, No. 12, pp (5318–5327).
92. Liu, B., Wang, L., Jin, Y. H., Tang, F. & Huang, D. X. (2005). "Improved particle swarm optimization combined with chaos", *Chaos, Solitons & Fractals*, vol.25, No. 5, pp (1261–1271).
93. Liu, L., Yang, H. & Lai, M., (2005). "Electricity price forecasting model based on chaos theory ", *International Power Engineering Conference* , UK, pp (1-5).
94. Liu, Y., Qiao, N., Zhao, C., Zhuang, J. & Tian, G. (2018). "Using the AR–SVR–CPSO hybrid model to forecast vibration signals in a high-speed train transmission system", *J Rail and Rapid transit*, vol.0 , No.0, pp (2-14).
95. Manurung, A.H., Suhartono, D., Hutahayan, B. & Halimawan, N. (2016). "Probability Bankruptcy Using Support Vector Regression", *Management Department, Binus Business School, Bina Nusantara University*.
96. McBride, N. (2005). "chaos theory as a model for interpreting information system in organizations", *info systems Journal*, vol.15, No.3, pp (233-254).
97. Mishra, S. & Patra, S. K. (2008). "Short Term Load Forecasting Using Neural Network Trained with Genetic Algorithm & Particle Swarm Optimization", *National Institute of Technology, First International Conference on Emerging Trends in Engineering and Technology*, pp (606-611).
98. mzadeh, R.E., Jampour, M. (2013) ."Chaotic Genetic Algorithm based on Lorenz Chaotic System for Optimization Problems", *Intelligent Systems and Applications*, vol. 5, No.5, pp(19-24).
99. Nasim, A.S. (2015). "chaos based cryptography and image encryption" , *Master's Thesis, University of Applied Sciences, Germany*.
100. Ojemakinde, B. T. (2006) "Support Vector Regression for Non-Stationary Time Series " , *Master's Thesis, University of Tennessee, Knoxville*.
101. Pai, P. F. (2006). "System reliability forecasting by support vector machines with genetic algorithms", *Mathematical and Computer Modelling*, vol. 43, No. (3-

- 4) , pp(262–274).
102. Pai, P.F., Hong, W.C. (2005). "Support vector machines with simulated annealing algorithms in electricity load forecasting", *Energy Conversion and Management*, vol. 46, No. 17, pp (2669–2688).
103. Perea, P. R., Ruiz, J. C., Chaparro, D.G., Venzor, J. A. P., Carreón, A.Q. & Rosiles J.G. (2013). "Support Vector Machines for Regression: A Succinct Review of Large-Scale and Linear Programming Formulations", *International Journal of Intelligence Science*, Vol. 3 No. 1, pp. (5-14).
104. Prakash, A., Kiwanis, N., Tiwari, M.K.& Cohen, Y. (2008). "Modified immune algorithm for job selection and operation allocation problem in flexible manufacturing system", *Advances in Engineering Software*, vol.39, pp (219–232).
105. Qiang, S. & Pu, Y. (2018). "Short-term power load forecasting based on support vector machine and particle swarm optimization", *Journal of Algorithms & Computational Technology*, Vol. 0, No. 0, pp (1-8).
106. Qin, J., Niu.D., Qiu.J.& Ji, L. (2015). "The Application of Joint Optimization Method using GA for Load Forecasting", *School of Economics and Management, North China Electric Power University*, pp (2442-2449).
107. Rui Zhang, & Yan-bin Wu. (2010). "Chaos Immune Algorithm and its application to function optimization". *University of Mining and Technology, Sixth International Conference on Natural Computation, China*, pp (2852-2855).
108. Sarhani ,M., El Afia, A. (2015). "Electric load forecasting using hybrid machine learning approach incorporating feature selection", *Mohammed-V University Rabat, Morocco*, pp (1-7).
109. Sharma, Y., Gupta, N.& Dayanand. (2015). "Genetic Algorithms-Overview, Limitations and Solutions", *Advances in Computer Science and Information Technology* Vol. 2, No.4, pp (329-333).
110. Shen, G., Huang, X. (2011). " Advanced Research on Electronic Commerce, Web Application, and Communication" , *International Conference, China, Part I*, pp (1-465).

111. Smola, A., Scholkop, B. (2004). "A tutorial on support vector regression", *Statistics and Computing*, vol.14, No. 3, pp (199–222).
112. Song, L. & Xu, X. (2010). "Flexible job shop scheduling problem solving based on genetic algorithm with chaotic local search", *South China University of Technology, China, Sixth International Conference on Natural Computation*, pp (2356-2360).
113. Taher, S. A. & Amooshahi, M. K. (2012). "New approach for optimal UPFC placement using hybrid immune algorithm in electric power systems", *International Journal of Electrical Power & Energy Systems*, vol. 43, No. 1, pp (899–909).
114. Toshev, A. (2019). "Particle Swarm Optimization and Tabu Search Hybrid Algorithm for Flexible Job Shop Scheduling Problem – Analysis of Test Results", *Institute of Information and Communication Technologies*, Vol.19, No.4, pp (26-44).
115. Wang, D., Tan, D. & Liu, L. (2017). "Particle swarm optimization algorithm: an overview", *Soft Computing*, vol. 22, No. 2, pp (387–408).
116. Wang, K., Kovacs, G., Wozny, M. & Fang, M. (2006). "Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management", *International Federation for Information Processing*, Vol. 207, pp (840-846).
117. Wang, L. Y. & Zhao, W. G. (2010). "Daily Discharge Forecasting Based on Support Vector Regression", *Advanced Materials Research*, vol.113-116, pp (386–389).
118. Wang, W., Xu, D., & Qiu, L. (2010). Support vector machine with chaotic genetic algorithms for annual runoff forecasting. *Sixth International Conference on Natural Computation*, pp (671-675)
119. Wang, Y. M., Zhang, Z., Zheng, J. B., & Tong, L. L. (2014). "Chaotic Genetic Algorithm for Mixed Integer Programming Problem", *Applied Mechanics and Materials*, China, vols. 651-653, pp(2273–2277).

120. Wang, Y., Wang, B. & Zhang, X. (2012). "A New Application of the Support Vector Regression on the Construction of Financial Conditions Index to CPI Prediction", *Procedia Computer Science*, vols. 9, pp (1263–1272).
121. Wendt, H. (2005). "Dipomarbeit Support Vector Machines for Regression Estimation and their Application to Chaotic Time Series Prediction", *Universität Wien Fakultät für Elektrotechnik und Informationstechnik*, Wien, pp (1-89).
122. Wenjun, W., Jiang, F. (2016). "Hybrid Genetic Algorithm and Support Vector Regression Performance in CNY Exchange Rate Prediction", *School of Statistics and Mathematics, Zhongnan University of Economics and Law*, China, pp (0136-0139).
123. Xu, Z., Gao, y. & Jin, y. (2014). "Application of an Optimized SVR Model of Machine Learning" , *International journal of Multimedia and Ubiquitous Engineering*, Vol.9, No.6, pp(67-80).
124. Xuemei, L., Lixing, D., Yan, L., Gang, X. & Jibin, L. (2010). "Hybrid Genetic Algorithm and Support Vector Regression in Cooling Load Prediction", *Third International Conference on Knowledge Discovery and Data Mining*, pp (527-531).
125. Yan, X. F., Chen, D. Z. & Hu, S. X. (2003). "Chaos-genetic algorithms for optimizing the operating conditions based on RBF-PLS model", *Computers & Chemical Engineering*, vol. 27, No. 10, pp (1393–1404).
126. Yousefi, P., Courtice, G., Naser, G. & Mohammadi, H. (2020). "Nonlinear Dynamic Modeling of Urban Water Consumption Using Chaotic Approach (Case Study: City of Kelowna)", *Water*, vol. 12, No. 3, pp (2-22).
127. Yuan, F.C. (2012). "Parameters Optimization Using Genetic Algorithms in Support Vector Regression for Sales Volume Forecasting", *Department of Information Management, Yuan Ze University*, 2012, vols. 3, pp (1480-1486).
128. Yuan, X., Yuan, Y. & Zhang, Y. (2002). "A hybrid chaotic genetic algorithm for short-term hydro system scheduling", *Mathematics and Computers in Simulation*, vol. 59, No. 4, pp (319–327).

129. Zemzami, M., Elhami.N., Itmi, M. & Hmina, M. (2019). "An evolutionary hybrid algorithm for complex optimization problems", International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8, No.2, pp (126-133).
130. Zhang, J. (2011). "Applied Informatics and Communication", International Conference, Suzhou University, Part III, No. 50, China, Springer Heidelberg Dordrecht London New York.
131. Zhang, W. Y., Hong, W. C., Dong, Y., Tsai, G., Sung, J. T. & Fan, G. (2012). "Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting", Energy, vol. 45, No. 1, pp(850–858).
132. Zhang, X., & Feng, T. (2016). "Chaotic bean optimization algorithm", Soft Computing, vol. 22, No. 1, pp (67–77).
133. Zhang, Z. & Ye, S. (2011). "Long Term Load Forecasting and Recommendations for China Based on Support Vector Regression", International Conference on Information Management, pp (597-602).
134. Zhu, T., Zheng, H. & Ma, Z. (2019). "A chaotic particle swarm optimization algorithm for solving optimal power system problem of electric vehicle", Advances in Mechanical Engineering, Vol.11, No. 3. pp (2-9).

مواقع الكترونية

- <https://www.educba.com/support-vector-regression/>
- <https://www.alukah.net/sharia/0/21969/#ixzz74H1yrMTa>

A decorative border of red starburst shapes surrounds the text. The border consists of a top row of 18 stars, a bottom row of 18 stars, and vertical columns of 18 stars on the left and right sides.

الاستنتاجات والتوصيات

أولاً: الاستنتاجات

- تم تقديم نموذج انحدار المتجه الداعم (SVR) مع ثلاث خوارزميات فوضوية تمثل: الخوارزمية الجينية الفوضوية (CGA)، خوارزمية المناعة الفوضوية (CIA)، خوارزمية سرب الطيور الفوضوية (CPSO)، لدراسة بيانات الحمل الكهربائي للمدة من 1980 الى 2019 للنتبؤ سنويا باستهلاك الطاقة الكهربائية في المنطقة الجنوبية. وتم التوصل الى الاستنتاجات الآتية:
1. يساعد الدمج بين نموذج انحدار المتجه الداعم (SVR) والخوارزميات الفوضوية في تحسين دقة التنبؤ وقد اعطى هذا النموذج تنبؤات جيدة وكفاءة وقريبة من القيم الفعلية لسلسلة الاستهلاك السنوي للطاقة الكهربائية.
 2. تم تهجين انحدار المتجه الداعم مع الخوارزميات التطورية الفوضوية لتطوير وتصميم اختيار أفضل المعلمات لانحدار المتجه الداعم (SVR).
 3. خلال المدة (1980 – 2019) لسلسلة بيانات الاستهلاك السنوي للطاقة الكهربائية اظهرت نتائج اختبار جذر الوحدة ان جميع متغيرات السلسلة الزمنية غير مستقرة عند المتوسط، وبعد اخذ الفرق الثاني (Defference = 2) استقرت السلسلة الزمنية لجميع متغيرات الدراسة للسلسلة الاصلية.
 4. أظهرت النتائج ان الخوارزمية الجينية الفوضوية (CGA) لها التفوق في التعامل مع عيوب الخوارزمية الجينية (GA) التي تعاني من الوقوع في الأمثلة المحلية (local optimization).
 5. بناء على اقل القيم لمعيار شوارتز (SBC) فضلا عن اصغر قيمة على وفق معايير (MSE،RMSE،MAE،MAPE) للمفاضلة بين النماذج المقترحة، اتضح النموذج (SVRCGA) انحدار المتجه الداعم (SVR) مع الخوارزمية الجينية الفوضوية (CGA) هو النموذج الافضل والملائم والأكثر كفاءة للنمذجة والتنبؤ بالحمل الكهربائي السنوي وذلك من خلال تفوقه على النماذج المقترحة الأخرى.

6. يقوم النموذج (SVRCGA) الذي يستند الى تقنية الخوارزمية الجينية الفوضوية (CGA) بحل مشكلة تحديد المعلمات كافة في نموذج انحدار المتجه الداعم (SVR)، ويعزز قدراته على التعميم.
7. تحسين دقة النموذج بشكل كبير والتنبؤ بالحمل الكهربائي في المنطقة الجنوبية للمدة (2020 – 2028) باستخدام النموذج الهجين (SVRCGA).
8. يعد تهجين تقنيات ذكية جديدة مختلفة بمساعدة خوارزميات تطويرية منفردة تكون أكثر ملاءمة للحصول على مستويات تنبؤ أكثر دقة.

ثانياً: التوصيات

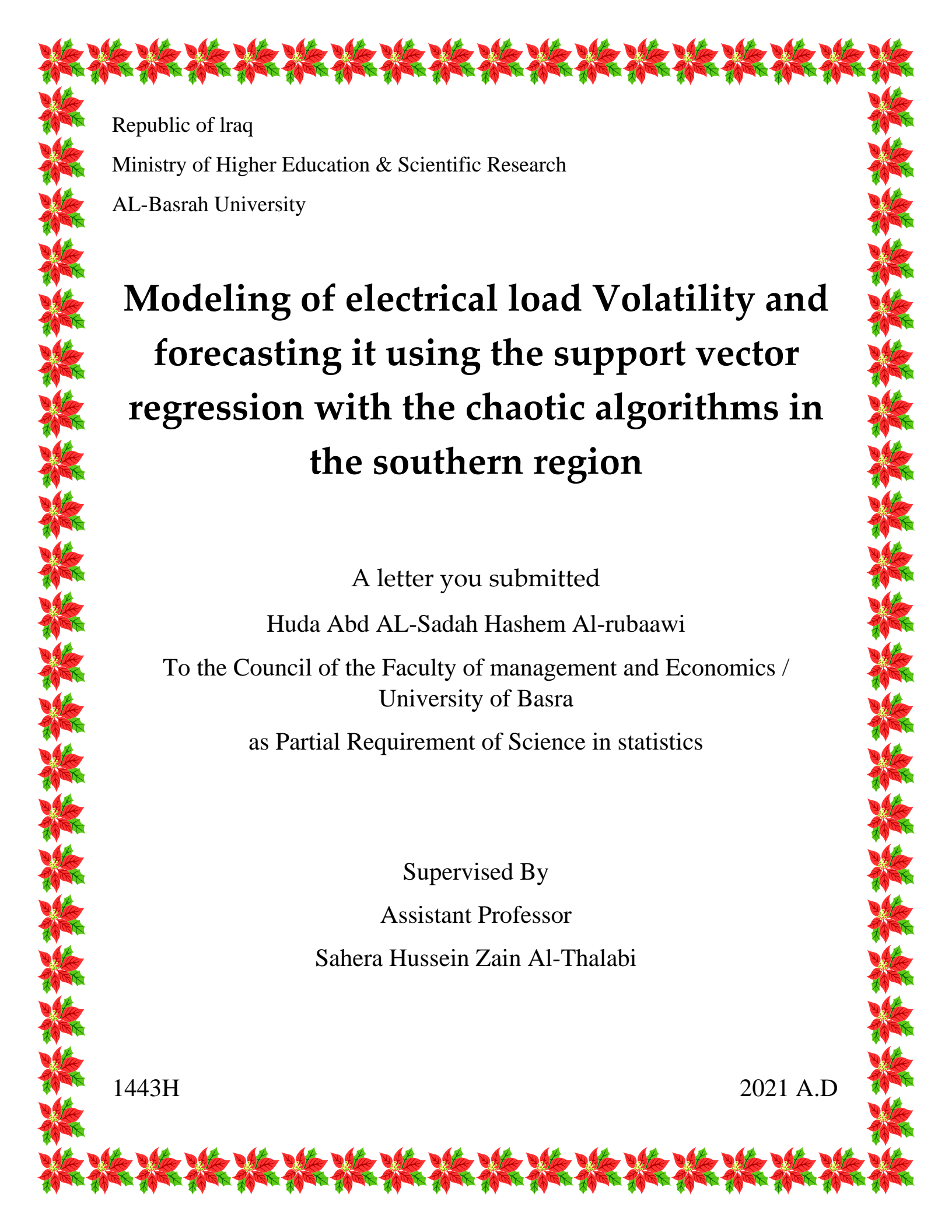
- بناء على النتائج التي تم الحصول عليها ندرج بعض التوصيات وكالاتي:
1. استخدام نتائج هذه الدراسة والاستفادة منها في عملية التنبؤ باستهلاك الطاقة الكهربائية في المنطقة الجنوبية للمدة (2020-2028) من قبل المعنيين والمسؤولين لصياغة الخطط والسياسيات المستقبلية المناسبة.
 2. استخدام النموذج (SVRCGA) الذي تم التوصل إليه في هذه الدراسة من قبل المعنيين للتحكم باستهلاك الطاقة وتحليل ودراسة هذه الظاهرة وفهم اتجاهاتها في المستقبل، اذ يكون لهذا النموذج اتجاهها واعداد للدراسة في المستقبل.
 3. استخدام خوارزميات التحسين المتقدمة الأخرى لتحديد معلمات انحدار المتجه الداعم (SVR) وتهجينها مع تقنيات وآليات أخرى او دمجها مع نموذج (SVR) والمقارنة فيما بينها لتحقيق تنبؤات أكثر دقة بالحمل الكهربائي.
 4. استخدام الخوارزميات الفوضوية لتقدير معلمات نموذج انحدار المتجه الداعم (SVR) ومقارنتها مع الخوارزميات والطرائق الهجينة الأخرى لغرض التوصل الى افضل نموذج للتنبؤ.

5. اجراء دراسات مناظرة في المستقبل على مستوى المحافظات عن تحسين نهج اختيار المعلمات المثالية استنادا الى الخوارزميات الفوضوية لزيادة الدقة والمقارنة فيما بينها، كما يمكن استكشاف قوه المعايير وحساسيتها في المستقبل.
6. الاستفادة من تقنية الخوارزميات الفوضوية لإيجاد الحلول المثلى في جميع المجالات التي تعني للقطاعات الأخرى من اجل الحصول على نتائج تنبؤ أفضل (اعلى موثوقية وأكثر دقة وكفاءة).

Abstract:

Load forecasting plays a key role in managing the power system and planning for scheduling power-generation capacity, and annual accurate forecasting of electrical loads can provide reliable guidance for power grid operation and planning. Moreover, accurate forecasting of electrical load has a larger and more important role in managing national energy systems strategy, since the load curve is non-linear, annual forecasts of electric power are a non-linear problem.

This study applies a new model for predicting electrical load by applying Support Vector Regression with chaotic hybrid algorithms to improve prediction performance, which resolves problems by improving all the parameters of the SVR. The stability of the SVR model depends on the choice of optimal parameters, A support vector regression is hybridized with three anarchic algorithms, chaotic genetic algorithm (CGA) ,chaotic Immune Algorithm (CIA), and chaotic the particle swarm Optimization (CPSO) algorithms to determine the optimal parameters for a supporting vector regression model (SVR). The models (SVRCGA) , (SVRCIA),(SVRCPSO) were then compared together to select the best model for forecasting electricity consumption for the period (2020-2028) in the southern region of Iraq. The results showed that the vector regression model supported with the anarchic genetic algorithm (SVRCGA) is superior, accurate and efficient than other prediction models based on statistical prediction accuracy criteria.



Republic of Iraq

Ministry of Higher Education & Scientific Research

AL-Basrah University

Modeling of electrical load Volatility and forecasting it using the support vector regression with the chaotic algorithms in the southern region

A letter you submitted

Huda Abd AL-Sadah Hashem Al-rubaawi

To the Council of the Faculty of management and Economics /
University of Basra

as Partial Requirement of Science in statistics

Supervised By

Assistant Professor

Sahera Hussein Zain Al-Thalabi

1443H

2021 A.D