

تعميم خوارزمية شعاع الدعم الآلي الضبابية ذات

نواة تشيبشيف لعدة أصناف

طالب الدراسات العليا: مرح قدورة كلية: العلوم – جامعة: دمشق

الدكتورة المشرفة: ريم القمحة

المخلص

نقدم في هذا البحث تعميماً لخوارزمية متجه الدعم الآلي الضبابية (Fuzzy Support Vector Machine) ذات نواة تشيبشيف المتعامدة (Orthogonal Chebyshev Kernel). تمكناً من تطوير الخوارزمية لتصبح قادرة على تصنيف أي عدد من الأصناف باستخدام تقنية فرق تسد (Divide and Conquer). نظراً لصعوبة حل مسألة التصنيف المتعدد قمنا بتبسيط المسألة وذلك بتقسيمها لعدة مسائل تصنيف ثنائي (Binary)، يتم في كل مسألة تحديد انتماء المدخلات لصنف معين أو عدم الانتماء له وذلك من خلال إيجاد دالة تصنيف خوارزمية متجه الدعم الآلي الضبابية لهذه المسألة تعرف هذه الطريقة بواحد مقابل الباقي (One-VS-Rest). بعد إيجاد دوال التصنيف لجميع المسائل يتم اختيار الانتماء لصنف المسألة التي أعطت دالة تصنيفها أعلى قيمة بالإضافة لإعطاء قيمة عضوية ضبابية تحدد درجة انتماء مدخل معين للصنف الناتج. قمنا بتوسيع مجال الدراسة في خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشيبشيف المتعامدة على $[-1, +1]$ ليشمل أي مجال مُعطى وبذلك تمكنا من إزالة عوائق تطبيق الخوارزمية على المجالات المختلفة.

الكلمات المفتاحية: شعاع الدعم الآلي الضبابي، نواة تشيبشيف المتعامدة، تقنية فرق تسد، تصنيف ثنائي، طريقة (One-VS-Rest)، قيمة عضوية ضبابية.

Generalization of the Fuzzy Support Vector Machine Algorithm with Kernel Chebyshev Orthogonal for Multi Classes

Dr. Reema Al-Kamha

Marah Qadoura

ABSTRACT

In this paper, we introduce a generalization of an Fuzzy Support Vector Machine Algorithm with Kernel Chebyshev Orthogonal algorithm. We were able to develop the algorithm to classify any number of classes using the Divide and Conquer technique. Due to the difficulty of solving the multiple classification problem, we have simplified the problem by dividing it into several binary classification problems. In each problem, the belonging of inputs to a certain class or not is determined by finding the classification function of FSVM algorithm. This method is known as One-VS-Rest method. After finding all classification function for all problems, the belonging will be chosen to the classes problem whose classification function gave maximum value, in addition to giving a fuzzy membership value that determines the degree of belonging of the specific input to the resultant classification.

We have extended the domain of study in the FSVM algorithm with kernel chebyshev orthogonal algorithm on $[-1, +1]$ to include any given domain, Thus we were able to overcoming problems of applying the algorithm to the different domains.

Keywords: Fuzzy Support Vector Machine, Orthogonal Chebyshev Kernel, Divide and Conquer technique, Binary Classification, (One-VS-Rest) method, Fuzzy Membership Value.

1- مقدمة:

تُعد خوارزمية متجه الدعم الآلي الضبابية (Fuzzy Support Vector Machine) من أهم خوارزميات التصنيف الضبابي في التعليم الآلي، حيث اكتسبت خوارزمية متجه الدعم الآلي الضبابية شهرة كبيرة في عدة مجالات، منها تصنيف الصور (Image Classification)، وتمييز الأرقام المكتوبة بخط اليد (Handwritten Digits Recognition)، وتمييز الأنماط (Pattern Recognition) والكثير من التطبيقات في المجالات الصناعية والمكتبية والزراعية والطبية.

تتلخص خوارزميات التصنيف الضبابي في البحث عن أفضل دالة تصنيف $\Phi: X \rightarrow Y$ يصنف مجموعة المدخلات $X = \{x_1, x_2, \dots, x_n\}$ لأحد الأصناف الموجودة مسبقاً في مجموعة المخرجات $Y = \{y_1, y_2, \dots, y_r\}$ بالإضافة لمقدار انتماء ضبابي S يُحدد درجة انتماء المدخل x_i $i = 1, \dots, n$; للصف y_j $j = 1, \dots, r$ [2].

تُستخدم خوارزمية متجه الدعم الآلي الضبابية في مسائل التصنيف الثنائي (Binary Classification)، حيث تقوم بتصنيف المدخلات لصفين فقط. الهدف الأساسي لعمل الخوارزمية هو إيجاد أفضل سطح (Separating Hyperplane) يفصل بين الصفين، حيث يوجد أكثر من مستوي فاصل يجب اختيار أفضل مستوي فاصل بين هذه المستويات [3].

تصنّف خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشيبيشيف النقاط التي تملك سلوك مُنتظم، أي تتوزع قيمها العُظمى والصغرى بشكل مُنتظم على المجال المدروس. من أهم ميزات هذه الخوارزمية أنها لا تحتاج لعدد كبير من متجهات الدعم مما يوفر الكثير من الوقت والذاكرة [4]، ولكنها تعاني من عدة مشاكل نذكر منها:

- يتم تصنيف النقاط لصفين فقط، وبالتالي لا يمكن استخدامها في المسائل التي تتطلب التصنيف لأكثر من صفين.
- ينحصر استخدام خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشيبيشيف المتعامدة على المجال $[-1, +1]$ ، وبالتالي لا تستطيع الخوارزمية حل مسائل التصنيف

على المجالات الأوسع.

درست العديد من الأبحاث تطوير خوارزمية شعاع الدعم الآلي لتتمكن من تصنيف النقاط لأكثر من صنفين وحل مشاكل التصنيف. قدم [5] طريقة لتبسيط مسألة تصنيف النقاط لعدة أصناف إلى مسائل شعاع دعم آلي جزئية متعددة كل منها تصنف النقاط لصنفين وذلك باستخدام دوال عضوية هرمية قطعية متعددة السطوح واختيار صنف المسألة التي نتج عنها أعلى قيمة. قدم [6] طريقة لتبسيط مسألة تصنيف النقاط في خوارزمية شعاع الدعم الآلي لعدة صفوف Multi-Class إلى مسائل جزئية متعددة كل منها تصنف النقاط إلى صف واحد Single-Class باستخدام بنية Kesler construction. قدم [7] أيضاً طريقة مبسطة جديدة لتصنيف النقاط لعدة صفوف تدعى الطريقة بـ Sim MSVM (Simplified Multi-Class SVM) استطاعت تخفيض حجم فضاء المسألة من بعدين إلى بعد واحد وتقليل خطأ التصنيف، بينت النتائج أن الخوارزمية تمكنت من تسريع عملية التدريب بشكل كبير مع الحفاظ على دقة التصنيف. قدم [8] طريقة لتصنيف النقاط لعدة أصناف ارتكزت هذه الطريقة على تقنية فرق تسد في خوارزمية شعاع الدعم الآلي سُميت الطريقة بـ DCSVM (Divide and Conquer Support Vector Machine) حيث درست النقاط المتناثرة في الفضاءات متعددة الأبعاد وقسمت مجموعة بيانات التدريب لمجموعات فرعية منفصلة لتسهّل عمل الخوارزمية.

2- هدف البحث:

نهدي في هذا البحث إلى تعميم التصنيف الثنائي (Binary Classification) في خوارزمية شعاع الدعم الآلي ذات نواة تشبيثيف إلى التصنيف لثلاث أصناف أو أكثر، وتوسيع مجال الدراسة من $[-1, +1]$ إلى أي مجال مُعطى، وذلك لحل المشاكل التي تعاني منها هذه الخوارزمية في التصنيف. تمكّننا من جعل الخوارزمية قادرة على التعامل مع أي مجال وبأي عدد من الأصناف وذلك بإرجاع المجال إلى $[-1, +1]$ وتدريب المدخلات على الانتماء إلى صنف من الأصناف المتعددة بالإضافة لمقدار انتماء ضبابي يعبر عن درجة انتماء الدخل لهذا الصنف.

3- أهم التعاريف والقوانين المستخدمة:

3-1- تعميم مجال الدراسة [1]:

يمكن إرجاع أي مجال $[a, b]$ حيث $a, b \in \mathbb{R}$ إلى المجال $[-1, +1]$ بإجراء التحويل التالي:

$$t = \frac{x - \frac{1}{2}(b+a)}{\frac{1}{2}(b-a)} \quad (1)$$

حيث $-1 \leq t \leq +1$ و $a \leq x \leq b$

بإجراء التحويل السابق على مدخلات الخوارزمية يتم إرجاع هذه المدخلات من أي مجال $[a, b]$ إلى المجال $[-1, +1]$ ، بالتالي يمكن تطبيق خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيشيف المتعامدة على أي مجال معطى من خلال إجراء تحويل الإرجاع للمجال $[-1, +1]$.

3-2- دالة الوزن [1]:

تكون الدالة $w(x)$ دالة وزن على المجال $[a, b]$ إذا وفقط إذا حقق ما يلي:

(1) الدالة w مستمر.

(2) $w(x)$ موجب تماماً.

(3) $\forall n \in \mathbb{N}; x \in [a, b] : \int_a^b w(x) |x^n| dx < +\infty$

3-3- الحدوديات المتعامدة [1]:

تكون مجموعة الدوال f_0, f_1, \dots, f_n متعامدة على المجال $[a, b]$ بالنسبة لدالة

الوزن $w(x)$ إذا تحقق ما يلي:

$$\int_a^b w(x) \cdot f_i(x) \cdot f_j(x) dx = \begin{cases} 0 & ; i \neq j \\ \text{عدد} & ; i = j \end{cases}$$

3-4- نواة تشبيشيف Chebyshev Kernel [1]:

تُعرف حدوديات تشبيشيف بأنها حدوديات متعامدة على المجال $[-1, +1]$

بدالة وزن $w(x) = \frac{1}{\sqrt{1-x^2}}$ كما يلي:

$$T_n(x) = \cos(n \cdot \arccos x) ; n \geq 0 , -1 \leq x \leq +1$$

$$n = 0 \Rightarrow T_0(x) = \cos(0 \cdot \arccos x) = 1$$

$$n = 1 \Rightarrow T_1(x) = \cos(1. \operatorname{arc} \cos x) = x$$

بالاستفادة من الخواص المثلثية:

$$\theta = \operatorname{arc} \cos x \Rightarrow \cos \theta = x$$

$$\Rightarrow T_n(\cos \theta) = \cos(n. \theta)$$

$$\left. \begin{aligned} T_{n+1}(\cos \theta) &= \cos((n+1). \theta) = \cos n\theta . \cos \theta - \sin n\theta . \sin \theta \\ T_{n-1}(\cos \theta) &= \cos((n-1). \theta) = \cos n\theta . \cos \theta + \sin n\theta . \sin \theta \end{aligned} \right\} \text{بالجمع}$$

$$T_{n+1}(\cos \theta) + T_{n-1}(\cos \theta) = 2 \cos n\theta . \cos \theta$$

بتعويض $\cos \theta = x$

$$T_{n+1}(x) + T_{n-1}(x) = 2T_n(x). x$$

يتم الحصول على العلاقة التكرارية التالية التي تربط بين حدوديات تشبيشيف:

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x) ; n \geq 1 \quad (2)$$

تُستخدم هذه العلاقة من أجل إيجاد حدوديات تقريب المربعات الصغرى من

الدرجة n والتي سنستخدمها كنواة لخوارزمية متجه الدعم الآلي، وتُكتب بالشكل:

$$\Phi(x) = c_0 T_0(x) + c_1 T_1(x) + \dots + c_n T_n(x) \quad (3)$$

لإيجاد حدوديات تقريب المربعات الصغرى للدالة $f(x)$ يتم تحديد المعاملات

c_0, c_1, \dots, c_n التي تجعل الخطأ أصغري، وهنا نميز حالتين:

(1) عندما تكون مجموعة النقاط مستمرة تُستخدم علاقة التعامد المستمر:

$$n, m \in N ; \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x). T_m(x) dx = \begin{cases} 0 ; n \neq m \\ \frac{\pi}{2} ; n = m \neq 0 \\ \pi ; n = m = 0 \end{cases}$$

تُحسب المعاملات بالشكل:

$$c_0 = \frac{1}{\pi} \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_0(x). f(x) dx$$

$$c_i = \frac{2}{\pi} \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_i(x). f(x) dx ; i = 1, \dots, n$$

(2) عندما تكون مجموعة النقاط منفصلة تُستخدم علاقة التعامد المنفصل:

$$n, m < N ; \sum_{k=1}^N T_n(x_k) \cdot T_m(x_k) = \begin{cases} 0 ; n \neq m \\ \frac{N}{2} ; n = m \neq 0 \\ N ; n = m = 0 \end{cases}$$

تُحسب المعاملات كالتالي:

$$c_0 = \frac{1}{N} \sum_{k=1}^N T_0(x_k) \cdot f(x_k)$$

$$c_i = \frac{2}{N} \sum_{k=1}^N T_i(x_k) \cdot f(x_k) ; i = 1, \dots, n$$

3-5- دالة الانتماء (العضوية) في المنطق الضبابي [9]:

تُكتب دالة الانتماء (العضوية) الضبابية للمدخل $x \in X$ بالشكل:

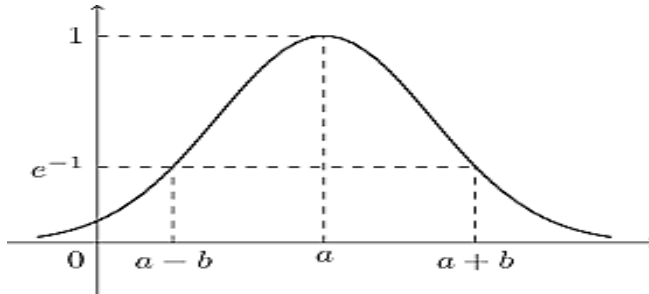
$$\mu: X \rightarrow [0,1]$$

حيث X المجموعة الشاملة التي تمثل كل القيم المحتملة للمدخل x .

تأخذ دوال العضوية عدة أشكال منها الدالة المثلثية (Triangular) ودالة شبه المنحرف (Trapezoidal) والدالة الجرسية أو الغاوسية (Gaussian) وهي الأكثر شهرةً لإمكانية تطبيقها على بيانات تتوزع بشكل عشوائي ومن أي طبيعة ممكنة، لذلك قمنا باستخدام الدالة الجرسية في هذا البحث، والتي تُعرف كما يلي:

$$G(x) = e^{-\left(\frac{x-a}{b}\right)^2} \quad (4)$$

يبين الشكل (1) الدالة الجرسية أو الغاوسية بالوسيطين a, b .



الشكل (1) دالة غاوس

4- خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيشيف المتعامدة [10]:

تُستخدم خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيشيف المتعامدة لتصنيف المدخلات إلى صنفين، كل مدخل $x_i \in \mathbb{R}^m ; i = 1, \dots, n$ ينتمي إلى أحد الأصناف $\{-1, +1\}$ في حالة $y_i = +1$ فإن الدخل ينتمي للصنف الأول، وفي حالة $y_i = -1$ فإن الدخل ينتمي للصنف الثاني، بالإضافة لقيمة ضبابية S_i تحدد درجة انتماء كل مدخل x_i للصنف y_i . تُدرَّب الخوارزمية مجموعة النقاط $(x_1, y_1, S_1), (x_2, y_2, S_2), \dots, (x_n, y_n, S_n)$ في حالة التصنيف الخطي، أما في حالة التصنيف غير الخطي تطبق دالة النواة على المدخلات $\Phi: x \rightarrow \Phi(x)$ ، فتصبح مجموعة نقاط التدريب كما يلي:

$$(\Phi(x_1), y_1, S_1), (\Phi(x_2), y_2, S_2), \dots, (\Phi(x_n), y_n, S_n))$$

يتم اختيار دالة النواة حسب طبيعة النقاط المدروسة. قمنا في هذا البحث بتصنيف النقاط التي تتوزع بانتظام على المجال $[-1, +1]$ باستخدام تابع نواة تشبيشيف المتعامدة بتطبيق العلاقة (3) على المدخلات، حيث تمكنا من خلال هذه النواة تحقيق نتائج أكثر دقة بمتجه دعم أقل [4].

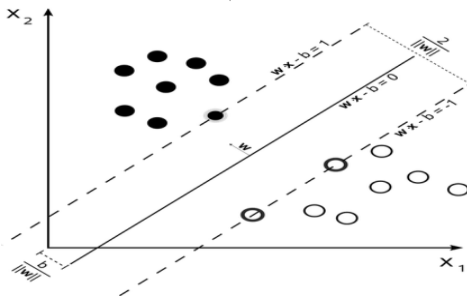
ترتكز فكرة خوارزمية متجه الدعم الآلي الضبابية في البحث عن أفضل فاصل بين النقاط يدعى الفاصل بالسطح الفائق، معرّف كما يلي:

$$w\Phi(x_i) + b = 0 \quad (5)$$

حيث $\Phi(x_i)$: ناتج تطبيق المدخل $x_i ; i = 1, \dots, n$ في العلاقة (3).

w و b : وسطاء دخل تقوم الخوارزمية بضبطها وأي تغيير فيها يؤثر على

مخرجات الخوارزمية حيث يتم البحث عن أفضل قيمة للثنائية (w, b) .



الشكل (2) تابع التصنيف الخطي لخوارزمية شعاع الدعم الآلي

تُصنّف نقاط المسألة لصنفين بحيث يفصل السطح الفائق بين الصنفين، أثناء عملية الفصل نأخذ أقرب نقطة إلى السطح الفائق من كل صنف من الصنفين ثم نبحث عن أكبر هامش Maximize the Margin، بحيث نقصد بالهامش بأنه المسافة بين المستوي الفائق وأقرب نقطة من الصنف والذي يمثل $\frac{b}{\|w\|}$ كما هو موضح في الشكل (2). لإيجاد أفضل سطح فائق علينا أن نبحث عن أفضل ثنائية لـ (w, b) باعتبار وجود خطأ في توضع النقاط تتشكل قيود الانتماء للأصناف كالتالي:

$$y_i(w \cdot \Phi(x_i) + b) \geq 1 - \epsilon_i \quad ; \quad i = 1, \dots, n$$

$$\epsilon_i \geq 0 \quad (6)$$

تؤول مسألة البحث عن أفضل هامش إلى المسألة الأمثلية التالية:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n S_i \epsilon_i \quad (7)$$

حيث C ثابت كفي.

تشكلت لدينا مسألة أمثلية (7) مقيدة بالقيود (6) ولحل هذه المسألة المقيدة من الضروري تحويلها لمسألة غير مقيدة، وذلك بضرب القيود بمضاريب لاغرانج α_i, β_i ، وإضافتها للمسألة:

$$L(w, b, \epsilon, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n S_i \epsilon_i$$

$$- \sum_{i=1}^n \alpha_i (y_i(w \cdot \Phi(x_i) + b) - 1 + \epsilon_i) - \sum_{i=1}^n \beta_i \epsilon_i$$

للبحث عن الحل الأمثل يتم الاشتقاق بالنسبة لـ w, b, ϵ_i .

$$\frac{\partial L(w, b, \epsilon, \alpha, \beta)}{\partial w} = w - \sum_{i=1}^n \alpha_i y_i \Phi(x_i) = 0 \quad (8)$$

$$\frac{\partial L(w, b, \epsilon, \alpha, \beta)}{\partial b} = - \sum_{i=1}^n \alpha_i y_i = 0 \quad (9)$$

$$\frac{\partial L(w, b, \epsilon, \alpha, \beta)}{\partial \epsilon_i} = S_i C - \alpha_i - \beta_i = 0 \quad (10)$$

بتعويض (8) و (9) و (10) في العلاقة (7) تصبح المسألة كالتالي:

$$\begin{aligned} \max W(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \sum_{i=1}^l y_i \alpha_i &= 0 \quad \text{ضمن الشروط} \\ 0 \leq \alpha_i &\leq S_i C, \quad i = 1, \dots, n \end{aligned} \quad (11)$$

يُطبق شرط Mercer [11] لحساب النواة $K(x, z)$ بين مدخلين متمايزين x, z من

مدخلات الخوارزمية كالتالي:

$$K(x, z) = \frac{\sum_{i=0}^n T_i(x) \cdot T_i(z)}{\sqrt{1 - x \cdot z}} \quad (12)$$

نلاحظ أن قيمة W مرتبطة بقيمة α . عند البحث عن قيمة α نستفيد من عبارة الخطأ ϵ ونطبق شروط كوهن توكر.

- عبارة الخطأ ϵ : هي الفرق بين الناتج المحسوب من (5) والناتج الفعلي y بالقيمة المطلقة:

$$\epsilon = |(w \cdot \Phi(x) + b) - y| \quad (13)$$

تُعوض العلاقة (8) في (13) فتتشكل عبارة الخطأ التالية:

$$\epsilon_i = \left| \left(\sum_{i=1}^n \alpha_i y_i \Phi(x_i) \cdot \Phi(x_i) + b \right) - y_i \right|$$

- شروط كوهن توكر **Kuhn-Tucker**: بالاستفادة من العلاقة (6) و (11) تُكتب الشروط:

$$(S C - \alpha) \epsilon = 0 \quad \text{ومنه} \quad \epsilon \geq 0 \quad \text{و} \quad S C - \alpha \geq 0$$

وهنا يتم تمييز حالتين:

- إما $0 < \alpha < S C$ أي $\epsilon = 0$ أي لا يوجد خطأ في تصنيف شعاع الدعم.
- أو $S C - \alpha = 0$ أي $S C = \alpha$ أي $\epsilon > 0$ ويوجد خطأ في التصنيف أو النقطة تقع على الخط الهامش. حيث α مرتبطة بـ S و C .

ضبط وسطاء الخوارزمية:

1) البحث عن القيم المثلى w :

للبحث عن أفضل قيم w يتم إيجاد القيمة المثلى لـ α وحدودها، لذلك يتم اختيار قيمتين مختلفتين α_1, α_2 من مجموعة قيم α .

• حدود α :

1- إذا كان α_1, α_2 من صنفين مختلفين، أي $y_1 \neq y_2$ تكون الحدود:

$$low = \max(0, \alpha_2 - \alpha_1)$$

$$high = \min(C(S_2 - S_1), C(S_2 - S_1) + \alpha_2 - \alpha_1)$$

2- إذا كان α_1, α_2 من نفس الصنف، أي $y_1 = y_2$ تكون الحدود:

$$low = \max(0, \alpha_2 + \alpha_1 - C(S_2 + S_1))$$

$$high = \min(C(S_2 + S_1), \alpha_2 + \alpha_1)$$

• إيجاد القيمة المثلى لـ α_1, α_2 :

$$\eta = 2 \Phi(x_1) \cdot \Phi(x_2) - \Phi(x_1) \cdot \Phi(x_1) - \Phi(x_2) \cdot \Phi(x_2)$$

$$\alpha_2 = -\alpha_2 + \frac{y_2(\epsilon_1 - \epsilon_2)}{\eta}$$

$$\alpha_2 \text{ clipped} = \begin{cases} high & ; \text{if } high \leq \alpha_2 \\ \alpha_2 & ; \text{if } low < \alpha_2 < high \\ low & ; \text{if } \alpha_2 \leq low \end{cases}$$

$$\alpha_1 = \alpha_1 + y_1 \cdot y_2 (\alpha_2 - \alpha_2 \text{ clipped})$$

وهكذا يتم ضبط حدود α وإيجاد القيمة المثلى لها، وبالتالي ضبط w .

3- البحث عن أفضل قيم b :

يتم اختيار قيمتين مختلفتين b_1, b_2 من مجموعة قيم b مع الأخطاء المرافقة

لها ϵ_1, ϵ_2 على الترتيب. وانطلاقاً من (5) و (8) نكتب:

$$b_1 = b - \epsilon_1 - \alpha_1 y_1 \Phi(x_1) \cdot \Phi(x_1) - \alpha_2 y_2 \Phi(x_1) \cdot \Phi(x_2)$$

$$b_2 = b - \epsilon_2 - \alpha_1 y_1 \Phi(x_1) \cdot \Phi(x_2) - \alpha_2 y_2 \Phi(x_2) \cdot \Phi(x_2)$$

$$b = \begin{cases} b_1 & ; \text{if } 0 < \alpha_1 < CS_1 \\ b_2 & ; \text{if } 0 < \alpha_2 < CS_2 \\ \frac{b_1 + b_2}{2} & ; \text{other than} \end{cases}$$

بعد ضبط الوسطاء w و b تنتهي مرحلة تدريب الخوارزمية. من أجل تصنيف أي دخل جديد نقوم بالاختبار التالي:

$$\text{sign}(w\Phi(x) + b) = \begin{cases} \text{انتماء للصنف الأول } y = +1 & \text{الناتج} \geq 0 \\ \text{انتماء للصنف الثاني } y = -1 & \text{الناتج} < 0 \end{cases} \quad (14)$$

نحصل في النهاية على صنف كل مدخل بالإضافة إلى قيمة S ، حيث قيمة S هي قيمة الانتماء الضبابية إلى هذا الصنف. وبذلك نكون قد صنفنا الدخل إلى الخرج المناسب له، مع قيمة عضوية ضبابية.

5- الخوارزمية المعدلة:

نقدم في هذه الفقرة تعميماً لخوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيشيف المتعامدة على المجال $[-1, +1]$ والتي تصنف المدخلات لصنفين، حيث تمكنا من تصنيف النقاط لعدة أصناف ومن أجل أي مجال معطى.

لحل أي مسألة تصنيف يتم بداية بإرجاع المجال $[a, b]$ إلى المجال $[-1, +1]$ بتطبيق تحويل الإرجاع من القانون (1)، ثم تبسيط مسألة تصنيف النقاط إلى عدة أصناف التي يصعب حلها إلى مسائل جزئية متعددة كل منها تصنف النقاط إلى صنفين باستخدام خوارزمية شعاع الدعم الآلي الضبابية، تكون نتيجة التصنيف في النهاية لصنف المسألة الذي يعطي أعلى قيمة.

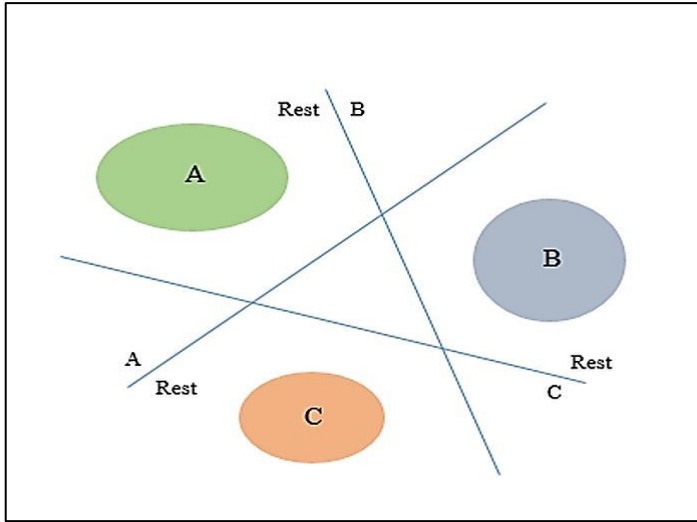
تُدرَّب الخوارزمية كل مدخل x_i (عدد المدخلات n وكل مدخل عبارة عن متجه له m مسقط أي $x_i \in \mathbb{R}^m$; $i = 1, \dots, n$) على الانتماء للصنف y_i (عدد الأصناف k التي قد تكون ثلاث أصناف أو أكثر أي $y_i \in \{3, 4, 5, \dots, k\}$; $i = 1, \dots, n$)، بالإضافة لقيمة ضبابية S_i تحدد درجة انتماء المدخل x_i للصنف y_i . بتطبيق دالة النواة على المدخلات $\Phi: x \rightarrow \Phi(x)$ ، تصبح مجموعة نقاط التدريب كما يلي:

$$(\Phi(x_1), y_1, S_1), (\Phi(x_2), y_2, S_2), \dots, (\Phi(x_n), y_n, S_n)$$

تُبسَّط مسألة تصنيف النقاط إلى عدة أصناف عن طريق تقسيمها إلى عدة مسائل جزئية كل منها تصنف النقاط إلى صنفين، تُعرف طريقة التقسيم هذه بتقنية فرق تسد

(Divide and Conquer)، ثم تُحل كل مسألة من هذه المسائل باستخدام طريقة واحد مقابل البقية (One-VS-Rest) [8].

يوضح الشكل (3) آلية عمل طريقة (One-VS-Rest) لتصنيف ثلاث أصناف A, B, C حيث يتم تبسيط مسألة التصنيف لثلاثة أصناف التي يصعب حلها بتقسيمها لثلاث مسائل تصنيف ثنائي وحل كل مسألة على حدى باستخدام خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيخيف بحيث تبحث كل مسألة عن أمثل سطح فائق يفصل بين أحد الأصناف وباقي الأصناف، مثلاً من أجل الصنف A نقوم بالبحث عن أمثل سطح فائق يفصل بين مدخلات التي تنتمي للصنف A وباقي المدخلات وتكون نتيجة دالة التصنيف الثنائي لمسألة متجه الدعم الآلي الضبابية المتعلقة بالصنف A إما الانتماء للصنف A أو عدم الانتماء للصنف A ، بنفس الطريقة يتم إيجاد دالة التصنيف للصنف B و للصنف C وفي النهاية نقرر انتماء المدخل لصنف المسألة التي تعطي أكبر قيمة في دالة التصنيف، حيث يوجد ثلاث توابع تصنيف في هذا المثال وهي دالة تصنيف للصنف A ودالة التصنيف للصنف B ودالة التصنيف للصنف C .



الشكل (3) يبين طريقة واحد مقابل الباقي لتصنيف ثلاث أصناف A, B, C

نفترض في خوارزمتنا أنه لدينا k صنف وبالتالي نحن بحاجة لإيجاد k نموذج تصنيف أي حل k مسألة تصنيف، كل صنف $r = 3, \dots, k$ يمثل مسألة يتم

تدريب الخوارزمية في هذه المسألة لصفين الأول ينتمي للصف r والثاني لا ينتمي للصف r ، وهكذا تحولت مسألة التصنيف المتعدد إلى تصنيف لصفين، نكرر هذه العملية على k صف وفي النهاية نحصل على k دالة تصنيف، نقرر انتماء المدخل لصف الدالة الذي يعطي أكبر قيمة.

خطوات الخوارزمية المعدلة:

من أجل كل صف r نبحث عن أفضل سطح فائق يحقق:

$$w^r \Phi(x_i) + b^r = 0 \quad (15)$$

حيث $\Phi(x_i)$: ناتج تطبيق المدخل x_i ; $i = 1, \dots, n$ في العلاقة (3).

w^r و b^r : وسطاء دخل الصف r تقوم الخوارزمية بضبطها وأي تغيير فيها

يؤثر على مخرجات الخوارزمية.

تهدف الخوارزمية للبحث عن أفضل سطح فائق، وبالتالي البحث عن أفضل ثنائية

لـ (w^r, b^r) وباعتبار وجود خطأ في توضع النقاط تشكل القيود التالية:

$$w^r \cdot \Phi(x_i) + b^r \geq 1 - \epsilon_i^r \quad ; \text{ if } y_i = r \quad (r \text{ ينتمي للصف } r)$$

$$w^r \cdot \Phi(x_i) + b^r \leq -1 + \epsilon_i^r \quad ; \text{ if } y_i \neq r \quad (r \text{ لا ينتمي للصف } r)$$

$$\epsilon_i^r \geq 0 \quad ; i = 1, \dots, n \quad (16)$$

تؤول مسألة الأمثلة لمسألة تصغير كالتالي:

$$\min \frac{1}{2} \|w^r\|^2 + C \sum_{i=1}^n S_i \epsilon_i^r \quad (17)$$

لحل هذه المسألة المقيدة من الضروري تحويلها لمسألة غير مقيدة، وذلك بضرب

القيود بمضاريب لاغرانج α_i, β_i ، وإضافتها للمسألة:

$$L(w^r, b^r, \epsilon^r, \alpha, \beta) = \frac{1}{2} \|w^r\|^2 + C \sum_{i=1}^n S_i \epsilon_i^r - \sum_{i=1}^n \alpha_i (y_i (w^r \cdot \Phi(x_i) + b^r) - 1 + \epsilon_i^r) - \sum_{i=1}^n \beta_i \epsilon_i^r$$

للبحث عن الحل الأمثل يجب الاشتقاق بالنسبة لـ w^r, b^r, ϵ_i^r .

$$\frac{\partial L(w^r, b^r, \epsilon^r, \alpha, \beta)}{\partial w^r} = w^r - \sum_{i=1}^n \alpha_i y_i \Phi(x_i) = 0 \quad (18)$$

$$\frac{\partial L(w^r, b^r, \epsilon^r, \alpha, \beta)}{\partial b^r} = - \sum_{i=1}^n \alpha_i y_i = 0 \quad (19)$$

$$\frac{\partial L(w^r, b^r, \epsilon^r, \alpha, \beta)}{\partial \epsilon_i^r} = S_i C - \alpha_i - \beta_i = 0 \quad (20)$$

بتعويض (18) و (19) و (20) في العلاقة (17) تصبح المسألة كالتالي:

$$\max W^r(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad \text{ضمن الشروط}$$

$$0 \leq \alpha_i \leq S_i C, \quad i = 1, \dots, n \quad (21)$$

استخدمنا شرط Mercer من العلاقة (12) لحساب النواة $K(x_i, x_j)$.

نقوم بإتباع خطوات الخوارزمية الأصلية لحساب أفضل w^r و b^r . بعد ضبط الوسطاء

w^r و b^r نحصل على أمثل سطح فائق يحقق:

$$w^r \Phi(x_i) + b^r = 0$$

نقوم بحساب الوسطاء w^r و b^r من أجل k صنف فنتتهي مرحلة تدريب الخوارزمية.

من أجل اختبار أي مدخل جديد نحسب قيمته في جميع الأسطح الفائقة وفي النهاية

ينتمي المدخل لصنف السطح الفائق الذي يعطي أكبر قيمة ممكنة بين الأسطح الفائقة

وتحسب كما يلي:

$$\arg \left(\max_r (w^r \Phi(x_i) + b^r) \right); r = 1, \dots, k \quad (22)$$

مثال عددي للخوارزمية المعدلة:

قمنا بأخذ مجموعة البيانات من (UCI Machine Learning Repository)

مستودع بيانات التعلم الآلي بجامعة كاليفورنيا في ايرفين [12]، حيث قسمنا المجموعة

إلى مجموعتين الأولى لمرحلة التدريب والثانية لمرحلة الاختبار.

أولاً: مرحلة التدريب:

يبين الجدول (1) بيانات تدريب الخوارزمية المؤلفة من 20 متجه دخل وكل متجه له 4 مساقط $x_i = (x_1, x_2, x_3, x_4)$ ويمثل y_i الخرج (الصنف) $y_i \in \{+1, +2, +3\}$.

الجدول (1) مجموعة بيانات التدريب للخوارزمية المعدلة

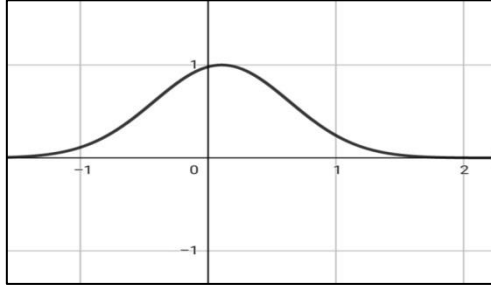
بيانات التدريب	x_1	x_2	x_3	x_4	y
1	1	1	2	3	+1
2	1	1	1	2	+1
3	1	1	1	3	+1
4	1	1	1	4	+1
5	1	1	1	5	+1
6	1	1	2	1	+1
7	1	1	2	2	+1
8	1	2	1	1	+2
9	1	5	1	1	+2
10	1	5	1	2	+2
11	1	5	1	3	+2
12	1	5	1	4	+2
13	1	5	2	1	+2
14	1	5	2	2	+2
15	1	2	1	2	+3
16	1	5	1	5	+3
17	2	1	2	1	+3
18	1	4	4	1	+3

19	1	4	1	4	+3
20	1	4	2	2	+3
21	1	3	3	1	+3

في البداية نقوم بإرجاع المجال من $[1,5]$ إلى $[-1,+1]$ باستخدام العلاقة (1). نعيين درجة انتماء المدخلات إلى الصنف الناتج وبما أنه لدينا ثلاثة أصناف بالتالي سنستخدم ثلاثة دوال انتماء $\mu: X \rightarrow [0,1]$.

(1) دالة الانتماء للصنف الأول وهي دالة غاوس كما يبين الشكل (4)، تعطى

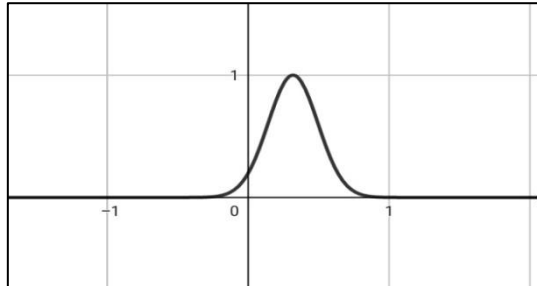
$$\mu = e^{-\left(\frac{x-0.1062}{-0.75}\right)^2} \quad \text{بالعلاقة:}$$



الشكل (4): دالة الانتماء الضبابي الغاوسية للصنف الأول.

(2) دالة الانتماء للصنف الثاني وهي دالة غاوس كما يبين الشكل (5)، تعطى

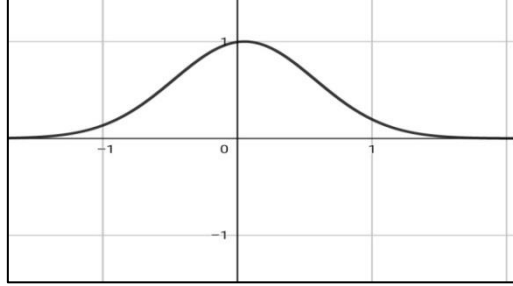
$$\mu = e^{-\left(\frac{x-0.3185}{0.25}\right)^2} \quad \text{بالعلاقة:}$$



الشكل (5): دالة الانتماء الضبابي الغاوسية للصنف الثاني.

(3) دالة الانتماء للصنف الثالث وهي دالة غاوس كما يبين الشكل (6)، تعطى

$$\mu = e^{-\left(\frac{x-0.05247}{-0.7428}\right)^2} \text{ بالعلاقة:}$$



الشكل (6): دالة الانتماء الضبابي الغاوسية للصنف الثالث.

نوجد حدوديات تقريب تشيبيشيف من الدرجة الرابعة من العلاقة (3) للمدخلات،

فنحصل على مجموعة نقاط التدريب كما يلي:

$$(\Phi(x_1), y_1, S_1), (\Phi(x_2), y_2, S_2), \dots, (\Phi(x_n), y_n, S_n)$$

حيث $\Phi(x_i)$: ناتج تطبيق المدخل x_i في العلاقة (3).

$$y_i : \text{صنف المدخل } x_i ; i = 1, \dots, n$$

$$S_i : \text{مقدار الانتماء الضبابي للمدخل } x_i \text{ إلى الصنف } y_i ; i = 1, \dots, n$$

نقوم بحساب الوسطاء w^r, b^r من أجل كل صنف r حيث $r \in \{+1, +2, +3\}$

$$\text{وذلك لنوجد أمثل سطح فائق } w^r \cdot \Phi(x) + b^r = 0$$

نُشكّل تابع لاغرانج ونشتقه ونوجد المعادلات (18), (19), (20) حتى نصل

للمسألة التالية:

$$\max W^r(\infty) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

حيث $K(x_i, x_j)$ نواة محسوبة من العلاقة (12).

نحتاج لبرنامج حاسوبي من أجل حل هذه المعادلة، لذلك قمنا باستخدام برنامج جافا.

الشكل (7) يبين إرجاع المجال من $[1, 5]$ إلى $[-1, +1]$ لعينة من النقاط ومقدار

الانتماء الضبابي لصنفها الناتج.

Support Vector	label	alpha	fuzzy
[-1.0, -1.0, -0.5, 0.0]	1.0	1.0	[0.11355950142 , 0.113559501427 , 0.5203284644360027 , 0.9801491157]
[-1.0, -1.0, -1.0, -0.5]	1.0	1.0	[0.11355950142 , 0.113559501427 , 0.1135595014272237 , 0.5203284644360027]
[-1.0, -1.0, -1.0, 0.0]	1.0	1.0	[0.11355950142 , 0.113559501427 , 0.1135595014272237 , 0.9801491157146593]
[-1.0, -0.5, -1.0, -1.0]	2.0	1.0	[8.31890145055 , 2.211894669965 , 8.3189014505555 , 8.318901450555513E-13]
[-1.0, 1.0, -1.0, -1.0]	2.0	1.0	[8.31890145055 , 5.925495854105 , 8.31890145055551 , 8.318901450555513E-13]
[-1.0, 1.0, -1.0, -0.5]	2.0	1.0	[8.31890145055 , 5.925495854105 , 8.31890145055551 , 2.2118946699658903E-5]
[-1.0, -0.5, -1.0, -0.5]	3.0	1.0	[0.13431158136 , 0.575112793532 , 0.134311581364 , 0.5751127935325844]
[-1.0, 1.0, -1.0, 1.0]	3.0	1.0	[0.13431158136 , 0.196478053007 , 0.13431158136 , 0.19647805300724447]
[-0.5, -1.0, -0.5, -1.0]	3.0	1.0	[0.57511279353 , 0.134311581364 , 0.5751127935 , 0.1343115813647233]

الشكل (7): يبين إرجاع النقاط للمجال $[-1, +1]$ ومقدار الانتماء الضبابي لصفها الناتج في برنامج الجافا.

وحصلنا على أمثل w و b ، للصف الأول والثاني والثالث كما يلي:

w	b
$\langle -6.690476190, 95.547619048, -6.476190476, 7.357142857 \rangle$	1
$\langle -5.642857143, 11.214285714, 10.809523810, 18.500000000 \rangle$	1
$\langle -16.785714286, 49.214285714, 22.666666667, 6.452380952 \rangle$	1

ثانياً: مرحلة الاختبار:

بعد انتهاء مرحلة التدريب نتأكد من صحة عمل الخوارزمية باختبار بيانات جديدة من خلال الدالة (22). لنختبر مجموعة النقاط التالية:

- $\{\{1,1,3,3\}, \{+1\}\}$

نتائج الاختبار في البرنامج الحاسوبي هو:

```
> to classify new point x & y (or exit):
1 1 3 3
Array2DRowRealMatrix{{-1.0,-1.0,0.0,0.0}}
class : 1|fuzzy [0.11355950142722376 0.11355950142722376 0.9801491157146593 0.9801491157146593]
```

تبين النتائج أن الخوارزمية قامت بإرجاع النقطة إلى المجال $[-1, +1]$ وأن النقطة تنتمي للصنف $+1$ بالإضافة لقيمتها العضوية الضبابية.

- $\{\{1,4,2,1\}, \{+2\}\}$

نتائج الاختبار في البرنامج الحاسوبي هو:

```
> to classify new point x & y (or exit):  
1 4 2 1  
Array2DRowRealMatrix{{-1.0,0.5,-0.5,-1.0}}  
class : 2|fuzzy [8.318901450555513E-13 0.5903285692842862 2.2118946699658903E-5 8.318901450555513E-13]
```

تبين النتائج أن الخوارزمية قامت بإرجاع النقطة إلى المجال $[-1, +1]$ وأن النقطة تنتمي للصنف $+2$ بالإضافة لقيمتها العضوية الضبابية.

- $\{\{4,3,3,4\}, \{+3\}\}$

نتائج الاختبار في البرنامج الحاسوبي هو:

```
> to classify new point x & y (or exit):  
4 3 3 4  
Array2DRowRealMatrix{{0.5,0.0,0.0,0.5}}  
class : 3|fuzzy [0.6955899848889683 0.9950226832510499 0.9950226832510499 0.6955899848889683]
```

تبين النتائج أن الخوارزمية قامت بإرجاع النقطة إلى المجال $[-1, +1]$ وأن النقطة تنتمي للصنف $+3$ بالإضافة لقيمتها العضوية الضبابية.

6- النتائج ومناقشتها:

يوضح المثال السابق أن خوارزمية متجه الدعم الآلي الضبابية المعدلة سخرت الخوارزمية الأصلية لحل كل مسألة من مسائل التصنيف الثلاثة وفي النهاية تم اختيار الانتماء لصنف المسألة التي تُعطي القيمة الأكبر، أي أن الخوارزمية المعدلة استطاعت تصنيف المدخلات لثلاثة أصناف أو أكثر. تبين النتائج مدى فعالية خوارزمتنا المعدلة، حيث كانت أشمل وأعم في التصنيف بالمقارنة مع الخوارزمية الأصلية التي تصنف المدخلات لصنفين فقط. وهكذا تكون الخوارزمية المعدلة قد حسنت من أداء خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشبيشيف المتعامدة.

إن توسيع المجال المدروس من $[-1, +1]$ إلى $[a, b]$; $a, b \in \mathbb{R}$ أضاف إمكانية التعامل مع أي مجال مُعطى، وبالتالي ساهمت الخوارزمية المعدلة بتحسين أداء الخوارزمية الأصلية بشكل واضح، وحلت بعض مشاكل الخوارزمية الأصلية على نحو فعال بإعطاء نتائج أكثر واقعية وأكثر فعالية.

7- الاستنتاجات والتوصيات:

قدمنا في هذا البحث تحسناً لخوارزمية متجه الدعم الآلي الضبابية ذات نواة تشييفيشيف المتعامدة. تمكنا أيضاً من توسيع المجال من $[-1, +1]$ إلى أي مجال مُعطى $[a, b]$. قمنا بتعميم التصنيف في خوارزمية متجه الدعم الآلي الضبابية ذات نواة تشييفيشيف لتصنف المدخلات إلى ثلاث أصناف أو أكثر بدلاً من التصنيف لصنفين فقط. كانت نتائج خوارزمتنا المعدلة أكثر واقعية وتطبيقها أكثر فعالية واستطاعت أن تزيل عائق المجال المحدود وتحل مسائل أشمل وأعم من الخوارزمية الأصلية. بالإمكان العمل مستقبلاً بتحسين الخوارزمية المعدلة باستخدام الأنظمة الموزعة لتقسيم المسائل المتعددة وحلها بكلفة أقل أو تشكيل نواة مركبة مع نواة تشييفيشيف للتعامل مع طبيعة جديدة وسلوك جديد للنقاط والاستفادة من هذا التعديل.

8- المراجع:

- [1] MATIT. B, 2014- Numerical Analysis (2). Damascus: Syria .Damascus University Publications, Faculty of Science. (in Arabic)
- [2] FU LIN. C, DE WANG. S, 2002- Fuzzy Support Vector Machines. IEEE Transactions on neural networks, 13(2),464-471.
- [3] WESTON. J, WATKINS. C, 1998- Multi-Class Support Vector Machines. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London.
- [4] MOGHADDAM. VH, HAMIDZADEH. J, 2016- New Hermite Orthogonal Polynomial Kernel And Combined Kernel In Support Vector Machine Classifier. Pattern Recognition, 60 ,921-935.
- [5] ABE. S, INOUE. T, 2002- Fuzzy Support Vector Machines For Multiclass Problems. ESANN, 113-118.
- [6] FRANC. V, HLAVAC. V, 2002- Multi-Class Support Vector Machine. Object recognition supported by user interaction for service robots, 2, 236-239.
- [7] WANG. Z, XUE. X, 2014- Multi-Class Support Vector Machine. Support Vector Machines Applications, 23-48.
- [8] DON. R, PRASANNA W. D, 2018- Multiclass Classification Using Support Vector Machines. Electronic Theses and Dissertation, 1845.
- [9] GEORGE J. KLIR, Bo. YUAN, 2008- Fuzzy Sets And Fuzzy Logic Theory And Applications. ISSN 0-13-101171-5.

- [10] AI-KAMHA. R, QADOURA. M, 2021- Fuzzification Of A Support Vector Machine Algorithm With Kernel Chebyshev Orthogonal. Journal of albaath University for Applied Sciences, 43.
- [11] JAFARZADEH. SZ, AMINIAN. M, EFATTI. S, 2013- A Set of New Kernel Function For Support Vector Machines: An Approach Based On Chebyshev Polynomials. ICCKE 2013, 412-416.
- [12] DUA. D, GRAFF. C, 2019- UCI Machine Learning Repository. From [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University Of California, School Of Information and Computer Science.
- [13] RAKOTOMALALA. R, 2017- SVM Support Vector Machine. Supervised Learning-Classification, 2, 48.
- [14] YE. N, SUN. R, LIU. Y, CAO. L, 2006- Support Vector Machine With Orthogonal Chebyshev Kernel. 18th International Conference on Pattern Recognition, (ICPR'06)2, 752-755.

