استخدام التعلم الآلي: الشبكات العصبونية وآلية متجه الدعم في تقدير وضع مريض سرطان الرئة

*دفاطمة شلاف

الكلمات المفتاحية: الشبكات العصبونية، التنبؤ، النمذجة، التعلم الآلي، آلية متجه الدعم، بيانات التدريب،

الملخص

يهدف البحث إلى بناء نماذج تعلم الآلي باستخدام الشبكات العصبونية وآلية متجه الدعم الـــــSVM ومقارنة النتائج بين كلا النوعين من حيث الدقة في تقدير وتنبؤ حالة الشـخص (١- مصـاب ٢- مهيأ ٣- سـليم) للإصـابة بسـرطان الرئة، وكانت العوامل المسـتقلة هي قيم (حمض البول – البوتاسيوم – الصوديوم –الارجنين)، حيث جمعت البيانات من مشفى البيروني في دمشق ومشـفى جامعة البعث التعليمي، تم جمع البيانات عن ٨٦ شـخص موزعين كما يلي: 50 مصـاب من مشـفى البيروني في دمشـق، و١٦ مهيأ (كلهم مدخنين) من مشـفى جامعة البعث و٢٠ سـليم من مشفى جامعة البعث.

في البداية تم بناء نموذج آلية متجه الدعم SVM وقســمت البيانات الى قســمين قســم التدريب 29 قيمة وقسـم الاختبار ٥٧ قيمة اختيرت بشـكل عشـوائي من بين القيم وحسـبت دقة بيانات الاختبار فكانت 100% ودقة بيانات التدريب ٨٦% ثم طبق النموذج على كل البيانات فكانت دقته 95.34% وهذا يؤكد لنا فعالية آلية متجه الدعم في الكشـــف عن حالة المريض بشكل مبكر. تم بناء نماذج شبكات عصبونية وتبين لنا من خلال المحاكاة أن افضل نموذج هو: NN(2,2) حيث تم تقســيم البيانات الى قســم تدريب 80% وقســم اختبار ٢٠% فكانت دقة الشـبكة بالنسـبة لبيانات التدريب 100% ولبيانات الاختبار ٩٤% وكانت الدقة بشـكل عام 98% وبالمقارنة مع نموذج آلية متجه الدعم SVM نجد أن هناك تفوق للشبكات العصبونية في تصنيف حالة المريض وننصح في استخدامها في تشخيص هذا النوع من الأمراض بيانات الاختبار.

*مدرس- قسم الإحصاء الرياضي- كلية العلوم -جامعة حمص.

Using Machine Learning: Neural Networks and Support Vector Machine in Estimating the Condition of Lung Cancer Patients

Fatima Shallaf*

Abstract

The aim of this research is to build machine learning models using neural networks and Support Vector Machine (SVM), and compare the results of both types in terms of accuracy in estimating and predicting a person's condition (1-affected, 2-predisposed, 3-healthy) for lung cancer. The independent variables were values of (uric acid, potassium, sodium, arginine). Data were collected from Al-Biruni Hospital in Damascus and Al-Baath University Teaching Hospital. Data was gathered from 86 individuals distributed as follows: 50 affected patients from Al-Biruni Hospital in Damascus, 16 predisposed individuals (all smokers) from Al-Baath University Hospital, and 20 healthy individuals from Al-Baath University Hospital.

Initially, an SVM model was built, and the data was divided into two parts: 29 values for training and 57 values for testing, chosen randomly. The accuracy of the test data was 100%, and the accuracy of the training data was 86%. The model was then applied to all the data, achieving an overall accuracy of 95.34%, confirming the effectiveness of the SVM in early detection of the patient's condition.

Neural network models were also built, and through simulation, it was found that the best model was: NN(2,2), where the data was split into 80% for training and 20% for testing. The network's accuracy for the training data was 100%, and for the test data, it was 94%, with an overall accuracy of 98%. In comparison with the SVM model, it was found that neural networks outperform in classifying the patient's condition, and thus, their use in diagnosing this type of disease is recommended

Keywords: Neural Networks, Prediction, Modeling, Machine Learning, Support Vector Machine (SVM), Training Data, Test Data.

*Teacher, Dept. of Mathematical Statistics, Faculty of Science, Homs University.

مقدمة (Introduction):

تُعد **تقنيات التعلم الآلي** اليوم جزءاً لا يتجزأ من الذكاء الاصطناعي(AI) ، وهو مجال يتطور بسرعة كبيرة ويشهد تقدماً هائلاً لقد بدأت رحلة الذكاء الاصطناعي رسمياً في عام ١٩٤٣ عندما قدّم كلٌ من **McCulloch**و **Walter Pitts**أول نموذج لشبكة عصبية، وهي محاولة لمحاكاة وظائف الدماغ البشري باستخدام أنظمة حسابية. ثم جاء **Alan Turing**في عام ١٩٥٠، حيث أسهم بشكل كبير في تطور هذا المجال عبر طرحه السؤال الشهير: "هل تستطيع الآلات التفكير؟"، وهو السؤال الذي فتح الباب لمزيد من الأبحاث حول إمكانية تطوير أنظمة ذكية تُحاكي القدرات البشرية. وفي عام ١٩٥٦، تم استخدام مصطلح "الذكاء الاصطناعي" لأول مرة خلال ندوة علمية عقدت في **معهد ماساتشوستس للتكنولوجيا(MIT)** ، وهي الندوة التي كانت بمثابة نقطة الانطلاق لهذا المجال الذي يشهد تقدمًا متسارعًا حتى يومنا هذا.

التعلم الآلي (Machine Learning)هو أحد الفروع الرئيسية للذكاء الاصطناعي، ويُعرّف بأنه القدرة على تمكين الآلات من "التعلم" من البيانات بدون الحاجة إلى برمجتها بشكل صريح لتنفيذ مهام محددة. بعبارة أخرى، يتم تدريب الآلات على فهم الأنماط واكتشاف العلاقات في البيانات لتحقيق نتائج دقيقة. ينقسم التعلم الآلي إلى نوعين رئيسيين:

١. **التعلم تحت الإشراف(Supervised Learning)** ، حيث يتم تزويد الآلة ببيانات مدخلة مصحوبة بالنتائج المطلوبة (المخرجات)، مما يمكنها من تعلم كيفية التنبؤ بالنتائج بناءً على المدخلات الجديدة.
٢. **التعلم غير الخاضع للإشراف(Unsupervised Learning)** ، حيث يتم إعطاء الآلة بيانات غير مصحوبة بالمخرجات، ويُطلب منها اكتشاف الأنماط أو العلاقات المخفية داخل البيانات.

في هذا البحث، سيتم الاعتماد على **أهم طرق التعلم الآلي**، وهما :**الشبكات العصبية (Neural Networks)**وآلية **دعم المتجهات .(Support Vector Machine – SVM)**

سنستعرض في هذا البحث شرحاً تفصيلياً لكل من هاتين الطريقتين، موضحين كيفية عملهما ودورهما في تقدير وتنبؤ حالة مرضى سرطان الرئة بناءً على البيانات الطبية المتاحة.

مشكلة البحث:

إن الكثير من الأمراض لا يتم اكتشافها الا في وقت متأخر وخاصة أمراض السرطان ومنها سرطان الرئة، وهذا يتسبب في فقدان الكثير من البشر لحياتهم رغم أنه كان بالإمكان انقاذهم من الموت.

أهمية البحث:

يعتبر اكتشاف مرض سرطان الرئة في مرحلة مبكرة من خلال تصنيف المريض كمهيأ للمرض مهماً جداً من أجل البدء بالعلاج قبل تطور المرض، ومن هنا تأتي أهمية البحث في استخدام الذكاء الاصطناعي والتعلم الآلي في محاولة الكشف عن المهيئين للمرض قبل فترة زمنية من استفحال المرض.

الهدف من البحث:

١- إيجاد فلتر ذكي لتصنيف المرضى لثلاث مجموعات (مصاب، سليم، مهيأ) وذلك بطريقتي الشبكات العصبونية، وآلية متجه الدعم.

٢- المقارنة بين طريقتي الشبكات العصبونية وآلية متجه الدعم، من حيث الدقة في التصنيف.

٣- إضافة طرق إحصائية ذكية لتشخيص سرطان الرئة في سوريا مما يكشف المرض مبكراً وبالتالي معالجته قبل تأزم وضع المريض.

١-خوارزمية آلية متجه الدعم (Support Vector Machine):[4][2][3]

إن خوارزمية آلية متجه الدعم (Support Vector Machine) والتي قد تختصر الى (اخصهراً) (SVM) من الطرائق الإحصائية التي تنتمي الى الذكاء الاصطناعي، حيث قُدمت لأول مرة من قبل العالم Vladimir Vapnik عام ١٩٩٢، وهي عبارة عن خوارزمية تعلّم عن طريق مشرف أو موجّه ، تُستعمل للتصنيف مُستندة إلى نظرية التعلّم

الإحصائية (Statistical Learning Theory).

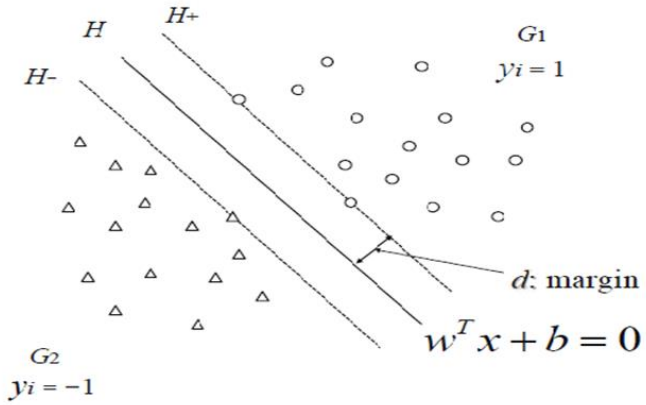
تَستخدم هذه التقنية مجموعة (جزء من البيانات) للتدريب تسمى بمتجهات الدعم (Support Vector) لتمثيل حد القرار (Decision Boundary) من خلال البحث عن أكبر هامش تبتعد عنه البيانات من الطرفين، وهو مستو فائق ذو هوامش أعظمية (Maximum Margin Hyperplane).[5]

ولتكن $X = (X_1, X_2, \dots, X_p)^T$ مجموعة من المتغيرات لبيانات تصنيف ثنائية ويراد تصنيفها في مجموعتين G_1 , G_2 ، وبأخذ مجموعة تدريب $x_1, x_2, \dots, x_n$ ولها نفس المتغيرات p ، الهدف هو العثور على مستو يصنف بيانات التدريب إلى مجموعتين G_1،G_2، فإذا لوحظ أنّ النقطة x_i تنتمي إلى المجموعة G_1 نضع $y_i = 1$ وإذا كانت تنتمي للمجموعة G_2 وضعنا $y_i = -1$، يمكننا تمثيل n من بيانات التدريب على النحو التالي:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n); y_i = \begin{cases} 1 & if \quad x_i \in G_1 \\ -1 & if \quad x_i \in G_2 \end{cases} \qquad (1)$$

وسنعرض فيما يلي آلية متجه الدعم في حالة البيانات قابلة للفصل خطياً وغير القابلة للفصل خطياً.

أولاً: آلية متجه الدعم في حالة بيانات قابلة للفصل خطياً Linear Separable SVM: [3][6][4]



الشكل (١) بيانات قابلة للفصل خطياً

نرسم مستقيمين متوازيين H_+ , H_- من خلال نقاط الحافة لكل مجموعة، نرسم حد القرار H وهو منتصف المسافة بين المستقيمين المرسومين عند نقاط الحافة وموازياً لهما، يمكن ملاحظة وجود أكثر من مستو لفصل البيانات إلا أنّنا نبحث عن المستو الذي تبتعد أو تقع عليه متجهات الدعم، بحيث يجب أن يكون أبعد ما يمكن عن نقاط المجموعتين.

تُصنف بيانات التدريب في مجموعتين G_1، G_2بالشكل التالي:

$$x_i \in G_1 \Leftrightarrow w^T x_i + b \geq 1 \qquad (2)$$
$$x_i \in G_2 \Leftrightarrow w^T x_i + b \leq -1$$

حيث: w الأوزان، b الانحياز.

وتعطى المسافة بين بيانات التدريب وحد القرار بالشكل التالي:

$$d = \frac{|w^T x + b|}{\|w\|} \; ; \; \|w\| = \sqrt{w_1^2 + w_2^2 + \cdots + w_p^2} \qquad (3)$$

بمعنى أن المستقيم الأول +H هو $w^T x_i + b = 1$

والمستقيم الثاني -H هو $w^T x_i + b = -1$

وبالتالي يكون حد القرار H هو: $w^T x_i + b = 0$

من أجل إيجاد الهامش بين المستقيمين +H و -H نقوم بحساب الفرق بين المستقيمن (إيجاد المسافة) أي إيجاد المسافة بين نقطة من كل حافة $wx_1 + b = 1$و $wx_2 + b = -1$

وبحساب الفرق نحصل على:

$$wx_1 - wx_2 = 2 \Rightarrow w(x_1 - x_2) = 2$$

وبعدها نحصل على:

$$\frac{w}{\|w\|}(x_1 - x_2) = \frac{2}{\|w\|}$$

وذلك للوصـــول على الهامش الذي يكون عمودي على المســتقيمات الثلاث وبالتالي تكون المســـافة هي $\frac{2}{\|w\|}$ والتي يجب أن تكون أكبر ما يمكن وعلينا الوصـــول إلى القيمة العظمى لها أو إيجاد القيمى الصغرى لـ $\frac{\|w\|^2}{2}$ كما يلي:

- **<u>تعلّم آلية متجه الدعم</u>**:[9][8][7] [4]

تهدف عملية التعلم الوصـــول لأكبر هامش تبتعد عنه البيانات من الطرفين، والتي تؤول إلى مسألة برمجة غير خطية مقيدة متمثلة بدالة الهدف:

$$f(w) = \min_{w} \frac{\|w\|^2}{2}$$

$$Subject\ to \qquad y_i(w^T x_i + b) \geq 1 \quad ; \quad i = 1,2,\dots,N$$

تحل هذه المســألة باســتخدام طريقة مضــاريب لاغرانج، وهي طريقة لإيجاد قيم النهايات الحدية المحلية لدالة لعدة متحولات خاضـــعة لواحد أو أكثر من القيود، نشـــكل دالة لاغرانج من أجل مسألة الأمثلية:

$$L(w, b, \lambda_1, \lambda_2, \dots \lambda_n) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \lambda_i (y_i(w^T x_i + b) - 1) \qquad (5)$$

حيث λ_i هي مضاريب لاغرانج.

نشتق دالة لاغرانج اشتقاق جزئي بالنسبة لـــ w,b وذلك لإيجاد الحد الأدنى لها ومن ثم نعطي هذه المشتقات الجزئية القيمة صفر، كما يلي:

$$\frac{\partial L(w, b, \lambda_1, \lambda_2, \dots \lambda_n)}{\partial \mathrm{w}} = 0 \Rightarrow w = \sum_{i=1}^{n} \lambda_i y_i x_i$$

$$\frac{\partial L(w, b, \lambda_1, \lambda_2, \dots \lambda_n)}{\partial \mathrm{b}} = 0 \Rightarrow \sum_{i=1}^{n} \lambda_i y_i = 0$$

نعوض في العلاقة (5) فنجد:

$$L_D(\lambda_1, \lambda_2, \dots, \lambda_n) = \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j \qquad (6)$$

وبالتالي نكون قد حصـــلنا على مســـألة مزدوجة بين w و λ_i ، تابع الهدف للمسألة المزدوجة يحتاج لإيجاد الحد الأعظمي:

$$Max\ L_D(\lambda_1, \lambda_2, \dots, \lambda_n) = \max \left\{ \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j \right\}$$

$$Subject\ to \quad \lambda_i \geq 0 \quad ; i = 1,2, \dots, n \qquad (7)$$

$$\sum_{i=1}^{n} \lambda_i y_i = 0$$

وهي مسـألة برمجة تربيعية (Quadratic Programming (QP)) بحلها نحصـل على:

متجـه الأوزان $\widehat{w} = \sum_{i=1}^{n} \hat{\lambda}_i y_i x_i$ والانحيـاز $\hat{b} = -\frac{1}{2}(\widehat{w}^T x_+ + \widehat{w}^T x_-)$

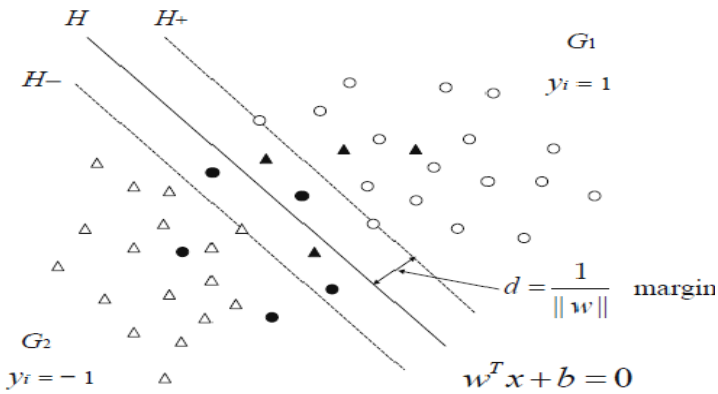
الأمثلي وحد قرار مع هوامش أعظمية $\widehat{w}^T x + \hat{b} = \sum_{i=1}^{n} \hat{\lambda}_i y_i x_i^T x + \hat{b} = 0$

وفقاً لذلك فإنّ قاعدة التصنيف للمجوعتين G_1،G_2 تكون كالتالي:

$$\widehat{w}^T x + \hat{b} = \sum_{i=1}^{n} y_i \hat{\lambda}_i x_i^T + \hat{b} = \begin{cases} \geq 0 \Rightarrow G_1 \\ < 0 \Rightarrow G_2 \end{cases} \qquad (8)$$

ثانياً: آلية متجه الدعم في حالة البيانات غير القابلة للفصـــل خطياً (Nonlinearly Separable SVM):[6][3]

تســـتخدم آلية متجه الدعم في هذه الحالة الهامش اللين (Soft Margin) الذي يسـمح لحد القرار بأن يحتمل أخطاء صـغيرة في بيانات التدريب كونها غير قابلة للفصـل خطياً، وبالتالي يعمل الهامش اللين على حل هذه المشكلة من خلال المقايضـة (Trade-Off) بين عرض الهامش وعدد أخطاء التدريب التي يرتكبها حد القرار من خلال تخفيض القيود الخطية بإدخال متحولات صورية غير سالبة ($\xi_i \geq 0$) .



الشكل (٢) بيانات غير قابلة للفصل خطياً

لتصبح القيود كمايلي:

$$y_i(w^T x_i + b) \geq 1 - \xi_i \quad ; i = 1,2,\dots,n \qquad (9)$$

حيث: $\xi_i = \max\left\{0, d - \frac{y_i(w^T x_i + b)}{\|w\|}\right\}$

- **تعلّم آلية متجه الدعم:[6][3][8]**

لتتجنب خوارزمية التعلّم تكبير هامش حد القرار الذي يؤدي للوقوع في أخطاء تصنيف الكثير من أمثلة التدريب، يتم تعديل حد القرار من خلال إعطاء قيم كبيرة لمتحولاته الصورية لتصبح مسألة البرمجة الخطية بالشكل:

$$f(x) = \min_{w,\xi}\left\{\frac{\|w\|^2}{2} + \alpha \sum_{i=1}^{n} \xi_i\right\} \qquad (10)$$

$$Subject\ to \quad y_i(w^T x_i + b) \geq 1 - \xi_i \quad ; \xi_i \geq 0, i = 1,2,\dots,n$$

حيث: $\alpha > 0$ تتحكم في المقايضة بين عرض الهامش والقيود.

نشكل دالة لاغرانج لمسألة الأمثلية المقيدة السابقة:

$$L(w,b,\xi,\lambda,\beta) = \qquad (11)$$
$$= \frac{1}{2}\|w\|^2 + \alpha \sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \lambda_i (y_i(w^T x_i + b) - 1 + \xi_i) - \sum_{i=1}^{n} \beta_i \xi_i$$

حيث: λ_i, β_i هي مضاريب لاغرانج.

باشتقاق دالة لاغرانج:

$$\frac{\partial L(w,b,\xi,\lambda,\beta)}{\partial w} = w - \sum_{i=1}^{n} \lambda_i y_i x_i = 0 \qquad (12)$$

$$\frac{\partial L(w,b,\xi,\lambda,\beta)}{\partial b} = -\sum_{i=1}^{n} \lambda_i y_i = 0 \qquad (13)$$

$$\frac{\partial L(w,b,\xi,\lambda,\beta)}{\partial \xi_i} = \alpha - \lambda_i - \beta_i = 0 \ ; i = 1,2,\dots n \quad (14)$$

نعوض في العلاقة (11) فنحصل على:

$$L_D(\lambda_1,\lambda_2,\dots,\lambda_n) = \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j \qquad (15)$$

من خلال العلاقة (15) نكون قد حصلنا على مسألة مزدوجة تابع الهدف لها بالشكل:

$$Max\ L_D(\lambda_1,\lambda_2,\dots,\lambda_n) = Max\left\{\sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j\right\}$$

$$Subject\ to\ \sum_{i=1}^{n} \lambda_i y_i = 0 \ \ ,0 \le \lambda_i \le \alpha \ ; i = 1,2,\dots,n \quad (16)$$

نلاحظ تطابق دالتي لاغرانج في حالتي البيانات القابلة للفصل خطياً وغير القابلة للفصل خطياً.

أمّا القيود على دالة لاغرانج في الحالتين مختلفة: ففي حالة البيانات القابلة للفصل خطياً كانت $\lambda_i \ge 0$ أمّا في حالة البيانات غير القابلة للفصل خطياً فإنّ λ_i يجب أن تحقق $0 \le \lambda_i \le \alpha$.

وبحل مسألة البرمجة التربيعية (16) نحصل على:

متجه الأوزان $\widehat{w} = \sum_{i=1}^{n} \hat{\lambda}_i y_i x_i$ والانحياز $\hat{b} = -\frac{1}{2}(\widehat{w}^T x_+ + \widehat{w}^T x_-)$

الأمثلين وحد قرار بهوامش أعظمية $\widehat{w}^T x + \hat{b} = \sum_{i=1}^{n} \hat{\lambda}_i y_i x_i^T x + \hat{b} = 0$

وفقاً لذلك فإنّ قاعدة التصنيف للمجوعتين G_1،G_2 تكون كالتالي:

$$\widehat{w}^T x + \hat{b} = \sum_{i=1}^{n} y_i \hat{\lambda}_i x_i^T + \hat{b} = \begin{cases} \ge 0 \Rightarrow G_1 \\ < 0 \Rightarrow G_2 \end{cases} \qquad (18)$$

٢- الشبكات العصبونية الاصطناعية Artificial Neural Networks: [9] [10] [11][12]

٢-١- مقدمة:

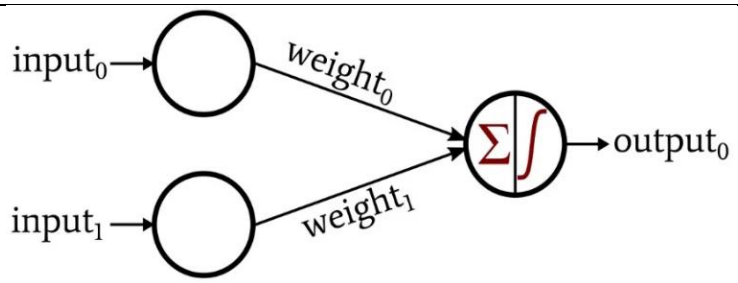
إن الشبكات العصبونية الاصطناعية هي محاكاة للطريقة التي يعمل بها الجهاز العصبي عند الانسان حيث وجد العلماء قابلية التعلم والتذكر والقدرة على التمييز واتخاذ القرارات، ومكونات الشبكة العصبونية الاصطناعية عبارة عن مجموعة من المعالجات التي تدعى بالعصبونات تمر الإشارات بين العصبونات عبر خطوط ربط ويرفق كل خط ربط بوزن معين(قيمة عددية) والذي يضرب مع الإشارات الداخلة إلى العصبون، ويطبق على كل عصبون تابع تفعيل ليحدد إشارة الخرج الناجمة عنه.

وبالتالي ستكون الشبكة العصبونية الاصطناعية مؤلفة من الأقسام التالية[13]:

١- طبقة الدخل input layer : هي طبقة الخلايا التي تستقبل إشارات الدخل من الوسط الخارجي

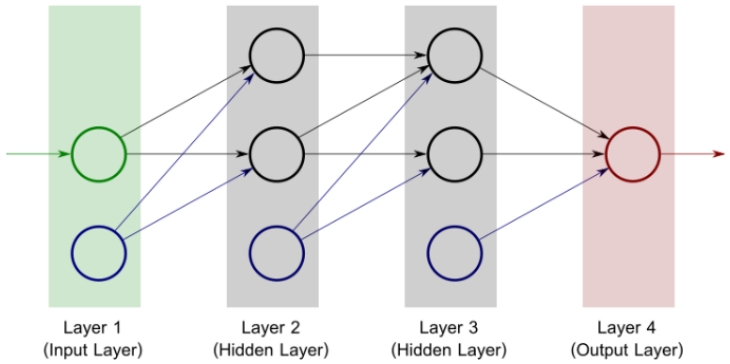
٢- طبقة الخرج output layer : هي طبقة الخلايا التي تعطي إشارات الخرج إلى الوسط الخارجي.

٣- الطبقة المخفية hidden layer : هي طبقة (طبقات) الخلايا التي تقع بين طبقتي الدخل والخرج وتقوم بعملية المعالجة، علما أنها تكون غير موجودة في الشبكة العصبونية البسيطة التي تكون مؤلفة من طبقة فقط (المدخلات وطبقة الخرج) كما في الشكل التالي



الشكل (3) شبكة عصبونية بسيطة وحيدة الطبقة.

أما في الشبكات العصبونية متعددة الطبقات يكون هناك طبقة مخفية واحدة أو أكثر من طبقة مخفية كما في الشكل التالي:



الشكل(4) شبكة عصبونية متعددة الطبقات.

وتدفق البيانات ضمن الشبكة العصبونية من طبقة الدخل الى الطبقات المخفية ثم طبقة الخرج وأن الشبكة العصبونية لا يمكن أن تتلقى من الخارج الا عن طريق هذه الطبقة (طبقة الدخل).

والجدير بالذكر أن الشبكة العصبونية يمكن أن تستوعب بيانات من عدد متغيرات أقل من عدد عصبونات الدخل بحيث نضع قيمة الصفر لعصبونات الدخل الاضافية وفي الحالة المعاكسة لا يمكن للشبكة العصبونية تلقي بيانات من عدد متغيرات أكبر من عدد عصبونات طبقة الدخل.

ويتم تحديد عدد المداخل وعدد عصبونات طبقة الخرج من المعطيات التي يتم تحليلها لذلك إذا كان ضمن المعطيات خمسة متغيرات لاستخدامها كمدخلات، ستبنى الشبكة العصبونية بخمسة مداخل، وبالمثل إذا كان هناك ثلاث متغيرات يجب دراستها كخرج للشبكة العصبونية تبنى الشبكة العصبونية بثلاث عصبونات في طبقة الخرج.

أما عدد عصبونات الطبقة المخفية لا يمكن تحديدها بسهولة حيث تُبنى بالتجريب، فلو كان عدد عصبونات الطبقة المخفية قليلاً فإن الشبكة العصبونية لن تتعلم، وان كان عدد عصبونات الطبقة المخفية كبيراً عندها ببساطة تميل الشبكة للحفظ، لذلك يتم تحديد عدد عصبونات الطبقة المخفية بين القلة التي تجعل الشبكة تتعلم وبين الكثرة التي تجعل الشبكة لا تحقظ معطيات الدخل.

(نبدأ ببناء الشبكة بعدد كبير نسبيا من عصبونات الطبقة المخفية وندرب الشبكة، هنا نعلم ان الشبكة تميل للحفظ وندرب الشبكة وثم نبدأ بتقليل عدد العصبونات بالتدريج إلى ان نحصل على أفضل النتائج مع الحفاظ على تدريب الشبكة العصبونية).

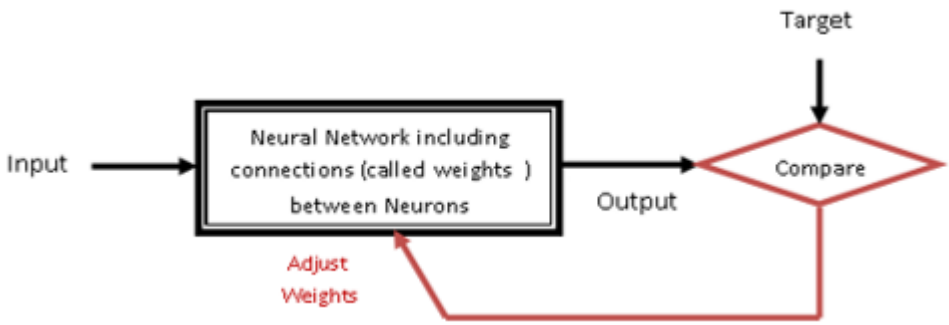
٢-٢- أنواع الشبكات العصبونية حسب نوع التدريب:[11][10][9]

حيث تعتبر عملية التدريب الأساس بالنسبة للشبكات العصبونية، لأنها عبارة عن عملية تستطيع الشبكة العصبونية من خلالها أن تتكيف مع مدخلات محددة من أجل الحصول على استجابات مطلوبة ويمكن تصنيف أساليب التدريب المختلفة للشبكات العصبونية إلى:

التدريب تحت الإشراف(المراقب) Supervised Training:

في التعليم المراقب (تحت الإشـــراف) تعطى مجموعة من بيانات التدريب المكونة من أزواج من(دخل – خرج) أي أن التدريب تحت الإشراف يحتاج إلى معلم أو مشرف يتمثل بشعاع الهدف ويتم التنبؤ بالخرج لدخل جديد وبعد إدخال معطيات التدريب تعالج الشبكة المدخلات وتقارن النتائج مع النتائج المرغوبة(الخرج المرغوب) وتحسـب الخطأ(الفرق بين الخرج الحالي الذي تم الحصـول عليه

من المدخلات وفق خوارزمية التدريب والخرج المرغوب)وتعاد (تنشـــر – تبث) الأخطاء من خلال الشـــبكة مما يؤدي لتحديث الأوزان التي تتحكم بالشـــبكة، تحدث هذه العملية عدة مرات طالما أن الأوزان تتعدل.



الشكل(5) طريقة التدريب بمعلم

التدريب غير المشرف عليه (غير المراقب) Unsupervised Training: [9][10][14][15][16]

التدريب بدون اشراف (بدون مراقبة)

لا تحتاج الشـبكات التي تسـتخدم هذا الأسـلوب من التدريب إلى معلم أي إلى شـعاع الهدف (المخرجات المرغوبة) فهي قادرة على التعلم والتصـــنيف للمدخلات المقدمة لها بشـكل ذاتي أي تقدم المدخلات للشـبكة بدون المخرجات المرغوبة، فهي نوع من أنواع شبكات التنظيم الذاتي، حيث تبني الشبكة العصبونية الاصطناعية أساليب التعليم على أســاس قدرتها على اكتشــاف الصــفات المميزة لما يعرض عليها من أشــكالٍ وأنســاق وقدرتها على تطوير تمثيل داخلي لهذه الأشكال وذلك دون معرفة مسبقة وبدون عرض أمثلة لما يجب عليها أن تُنتجه وذلك على عكس المبدأ المتبع في أســـلوب التعليم بوســاطة معلم، ويوجد عدة أنواع للشــبكات ذات خوارزميات التعلم بدون مشــرف نذكر منها:

١- شبكة الخريطة ذاتية التنظيم self_Organizing Mapping: أو شبكة Kohonen وهي نوع من شبكات التعلم التنافسي غير المشرف عليه وسميت بهذه التسمية بسبب التنافس بين العقد لتصبح نشطة وتدعى العقدة الرابحة حيث تقوم بإجراء تعديل على شعاع الوزن، وتنقل جزء من التعلم للعقد المجاورة لها.[15][14]

٢- الغاز العصبوني Neural Gas: تستخدم أيضاً في التدريب التنافسي وتطبق في مجالات التنبؤ، وتعتبر أكثر فعالية من شبكة الخريطة ذاتية التنظيم في مهام تنقيب المعطيات، حيث تقوم باكتشاف الطبوغرافية بنفسها وتعمل على ايجاد أفضل تمثيل لهيكلية المعطيات وإعطاء أفضل نتائج للتنقيب.[15][14]

٣- الغاز العصبوني المتنامي Growing Neural Gas: يشبه العنقدة التي تقوم على تجميع المعطيات (البيانات) في فضاء R^2 في عدد محدد مسبقاً من العناقيد حيث أن كل النقاط الموجودة في العناقيد حيث أن كل النقاط الموجودة ضمن العنقود متقاربة بالخواص فيما بينها أكثر من النقاط الموجودة في العناقيد الاخرى أما في الغاز العصبوني المتنامي يتم توزيع المدخلات ضمن R^n ويشكل بشكل متزايد شبكة من العقد حيث تمثل كل عقدة من الشبكة بنقطة من R^n، وفي حال تغير المدخلات فإن الغاز العصبوني قادر على تحريك العقد لتغطي التوزيع الجديد للدخل، أي لا يتم تحديد عدد العقد ولكنه يتم زيادتها بشكل متنامي أثناء التنفيذ ويم ايقاف إدراج العقد الجديدة عند قيام المستخدم بتعريف تابع الأداء أو في حال الوصول الى الحد الأقصى لحجم الشبكة، وهكذا يكون لكل عقدة ثلاث عناصر هي الأوزان المعرفة على R^n، والخطأ المتراكم، ومجموعة الحواف التي تعرف جيران العقدة.[14]

٢-٣- أنواع الشبكات حسب البنية:[13]

١- الشبكات العصبونية بتغذية أمامية feedforward networks

٢- الشبكات العصبونية العكسية Recurrent Neural Networks

الشبكات العصبونية بتغذية أمامية feed forward networks:[13]

تنتشر هنا المدخلات من وحدات الدخل الى الطبقات المخفية فطبقة الخرج عن طريق وصلات مقترنة بأوزان كما يلي:

المدخلات inputs: تستقبل الاشارات الواردة من المحيط الخارجي وتمثل بالشعاع $(x_1, x_2, \dots, x_n)$ حيث يتم نقلها إلى جسم الخلية بنسب مختلفة.

الأوزان weights: قيم عددية قابلة للزيادة والنقصان وفق معادلات رياضية محددة ويرمز لها بـ$(w_1, w_2, \dots, w_n)$، حيث يتم اعطائها قيم ابتدائية في الخطوة الأولى للتدريب ثم يتم تعديل هذه القيم حتى الوصول الى أصغر قيمة ممكنة للخطأ الذي يتمثل بالفرق بين الخرج الفعلي للشبكة العصبونية والخرج المرغوب.

تابع التجميع summation function: يقوم بجمع المدخلات الموزونة weighted function وفق العلاقة:

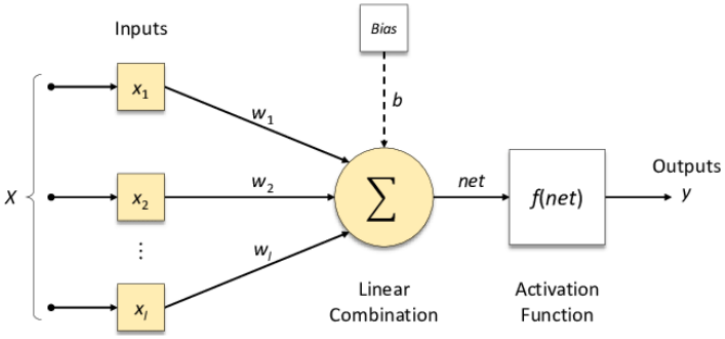
$$yin = \sum_{i=1}^{i=n} w_i x_i = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \quad (19)$$

تابع التفعيل (التنشيط)activation function : يقوم بتحويل حاصل جمع المدخلات الموزونة إلى إشارة رياضية محددة حسب طبيعة المشكلة لينتج الخرج النهائي للخلية العصبية كما يلي:

$$\boldsymbol{yin} = \sum_{i=1}^{n} w_i x_i$$

فيكون خرج الشبكة العصبونية بتطبيق تابع التفعيل على $\boldsymbol{yin}$ كما في العلاقة التالية:

$$output = f(yin) = f\left(\sum_{i=1}^{i=n} w_i x_i\right) \quad (20)$$



الشكل (6) الشكل التفصيلي للشبكة العصبونية بتابع التجميع والتفعيل.

أنواع تابع التفعيل:

١. تابع التفعيل الخطي
٢. تابع الخطوة الثنائي (ذو عتبة **θ**)
٣. تابع الاشارة (ذو عتبة **θ**)
٤. تابع سيغمويد (Sigmoid)
٥. تابع سيغمويد الثنائي(Sigmoid)

٣- التطبيق العملي:

تم قياس المشعرات الأربعة وهي الأرجنين وحمض البول والبوتاسيوم والصوديوم وجمعت البيانات عن ٨٦ شخص موزعين كمايلي: 50 مصاب، ١٦ مهيأ (كلهم مدخنين)، ٢٠ سليم.

أولاً: تطبيق آلية متجه الدعم:

حيث تم بناء نموذج التعلم الآلي SVM وقسمت البيانات الى قسمين قسم التدريب 29 قيمة كما يلي:

الجدول (1) بيانات التدريب لآلية متجه الدعم

رقم المريض	Y	Uricacid	Na	K	Arginine
3	1	11	145.3	3.69	12.6
6	1	4.4	136.7	5.22	29.7
14	1	6.4	133.3	3.57	21
16	1	1.8	134.3	3.58	11.9
19	1	3.8	139.9	4.29	14.8
20	1	5.4	138.7	4.45	24.5
25	1	2.8	144.5	4.32	9.1
26	1	4.4	138.6	3.94	13.6
28	1	4.8	136.7	4.44	23.7
30	1	5	142.1	5.05	22.9
43	1	2.9	144	4.64	12.4
45	1	5.1	138.2	4.5	18
49	1	3.7	142	4.93	16.3
51	2	4.1	138	4.2	20.4
52	2	3.8	136.2	4.4	55.1
55	2	5.1	152	4.3	47.9
57	2	3.3	150.1	5.3	29.7
58	2	5.2	139	4.1	31.4
59	2	4.1	148.2	5.1	28.2
60	2	3.3	136	4.8	33.1
61	2	5.8	145.4	4.1	21
62	2	4.1	132.7	3.9	44.7
63	2	6	139	4.6	27.5
65	2	2.9	142.5	4.8	37.1
66	2	3.8	137.6	4.6	21.9
67	3	5.4	137.1	4.81	68.3
77	3	3.1	136.6	5	82.9
82	3	4.8	139.1	3.64	76.8
86	3	3.7	142.2	3.84	84

أما بيانات الاختبار فكانت 57 قيمة بالشكل التالي:

الجدول (2) بيانات الاختبار لآلية متجه الدعم

رقم المريض	y	Uricacid	Na	K	Arginine
1	1	3.6	127.5	4.85	1.2
2	1	3.2	137.9	4.18	9.3
4	1	4	142.5	4.18	5.7
5	1	5.4	136.5	4.46	11.5
7	1	6.3	137.6	4.08	1
8	1	4.4	139.2	4.15	1.3
9	1	4.6	139.6	4.09	9.7
10	1	5.3	132.8	4.9	4.3
11	1	3.7	132.8	4.93	5.9
12	1	5.6	140.7	4.78	16.8
13	1	2.1	136.9	3.96	9.1
15	1	4	134.5	3.74	9.7
17	1	3.8	140.2	5.38	4.2
18	1	4.7	135.3	4.72	17.4
21	1	6.1	139.3	6.37	27.4
22	1	6.2	138.6	6.65	9.7
23	1	4	137.5	5.36	2.7
24	1	4.9	134.4	7.7	8
27	1	3.3	138	4.25	2
29	1	3	139.4	4.4	3.9
31	1	3	139.6	5.03	14.2
32	1	3.7	133.2	5.3	3.7
33	1	4.1	134.6	4.34	17.4
34	1	4.3	132.4	5.13	14.1
35	1	3.1	133.1	4.9	9.1
36	1	4.2	132.8	4.22	12
37	1	5	131	5.1	14.6
38	1	3.4	134.7	4.36	8.6
39	1	2.7	140.1	4.82	7
40	1	5.3	139.2	5.1	10.5
41	1	5	135.6	5	11.3
42	1	3.6	132.2	5.22	16
44	1	6	132.6	5.12	9.3

46	1	6.1	142.1	4.92	13.3
47	1	2.6	139.1	5.02	10.1
48	1	4.8	134.2	5.44	20
50	1	2.8	140.1	5.13	9.1
53	2	3.4	142.7	4.3	50
54	2	4.6	151	4.7	27.6
56	2	4.6	148.6	4.8	32
64	2	5.2	156	5.2	28
68	3	6.9	139.5	3.9	75
69	3	3.4	140.1	5.02	102.1
70	3	3.9	143	3.99	88
71	3	5.1	142.2	4.16	113
72	3	4.2	144.1	4.08	90.6
73	3	6.5	141	4.9	103
74	3	5	138.1	3.78	84.2
75	3	3.9	140.2	3.6	110
76	3	4.6	139	4.25	130.5
78	3	5.3	144.7	4.82	98
79	3	4	139.5	3.84	114.7
80	3	2.8	142.6	4.16	132
81	3	4.1	143	4.55	94.7
83	3	5.1	143.7	5.08	108.2
84	3	6.2	140	4.76	120.5
85	3	4.2	138.2	4.46	116

بلغت دقة التنبؤ في بيانات الاختبار الى ١٠٠ % وهو شرط أساسي ليكون النموذج قابل للتطبيق بينما كانت دقة بيانات التدريب ٨٦,٢١% وبعد تطبيق النموذج على كل البيانات كانت الدقة ٩٥,٣٥%

وكانت نتائج بيانات التدريب كما في الجداول التالي:

نسبة الخطأ	3	2	1	y
0	0	0	13	١
4/29	0	8	4	٢
0	4	0	0	٣
0.137931	نسبة الخطأ الكلي			

وكانت نتائج بيانات الاختبار:

نسبة الخطأ	3	2	1	y
0	0	0	٣٧	١
0	0	٤	٠	٢
0	١٦	0	0	٣
0	نسبة الخطأ الكلي			

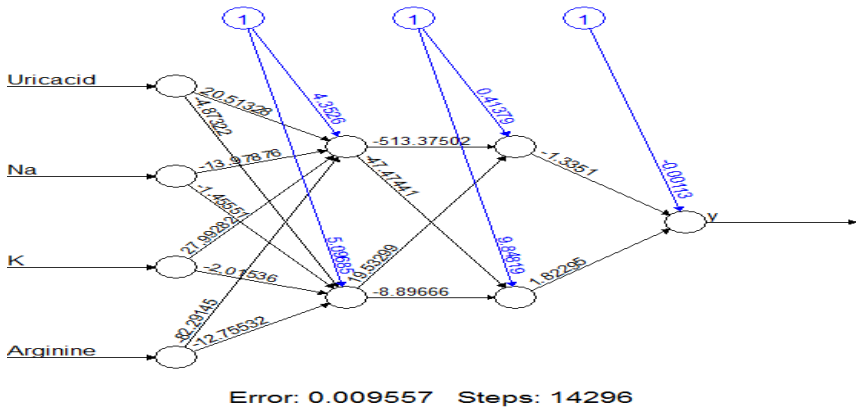
والنتائج النهائية لكل البيانات:

نسبة الخطأ	3	2	1	y
0	0	0	٥٠	١
4/86	0	١٢	4	٢
0	٢٠	0	0	٣
0.046512	نسبة الخطأ الكلي			

وهذا يؤكد فعالية آلية متجه الدعم في الكشف عن حالة المريض بشكل مبكر.

ثانياً: تطبيق الشبكات العصبونية:

تم بناء نماذج شبكات عصبونية وتبين لنا من خلال المحاكاة أن افضل نموذج هو: NN(2,2) حيث تم تقسيم البيانات الى قسم تدريب 80% وقسم اختبار ٢٠% وبنفس أسلوب الدراسة في آلية متجه الدعم كانت دقة الشبكة بالنسبة لبيانات التدريب 100% ولبيانات الاختبار ٩٤% وكانت الدقة بشكل عام 98% ويمكن تمثيل الشبكة بالشكل التالي:



الشكل (7) هيكل الشبكة العصبونية المدربة على البيانات

وبالمقارنة مع نموذج آلية متجه الدعم SVM نجد أن هناك تفوق للشبكات العصبونية في تصنيف حالة المريض وننصح في استخدامها في تشخيص هذا النوع من الأمراض.

النتائج:

١- حصلنا على دقة تنبؤ تصنيف سرطان الرئة باستخدام آلية متجه الدعم بلغت ٩٦%

٢- حصلنا على دقة في تصنيف سرطان الرئة باستخدام الشبكات العصبونية باغت (بلغت)٩٨%

٣- تفوقت الشبكات العصبونية على طريقة متجه الدعم في تصنيف حالات مرض سرطان الرئة

التوصيات:

١- استخدام طرق التعلم الآلي مثل الشبكات ومتجه الدعم في الكشف عن أمراض أخرى

٢- استخدام طرق العنقدة والانحدار اللوجستي واللوجستي الضبابي في الكشف عن الامراض.

المراجع:

[1]Mohammed Mohssen, Bashier Eihab Bashier,(2017),_ ***Machine Learning: Algorithms and Applications***_ by Taylor & Francis Group, LLC

[2] Suicheng Gu, YingTan , XinguiHe (2010),_ ***Discriminant analysis via support vectors.,*** *Elsevier*.

[٣] SADANORI K., -**Introduction to Multivariate Analysis**. Chuo University, Tokyo, Japan, 299.

[4] حباق هبة،2018 _ **مقارنة بين تحليل التمايز الخطي لفيشر /Fisher-LDA/ وآلية متجه الدعم /SVM/ للتصنيف الثنائي بين بيانات زيت الذرة وزيت دوار الشمس**_ جامعة حلب، سوريا، ١٢٥

[5] أبوزلام أحمد، ٢٠١٣–**دراســة بعض طرائق الشــبكات العصــبية وآلة متجه الدعم والطرائق الإحصائية في مسائل التصنيف لدعم القرار الطبي.** جامعة حلب، سوريا، ٩٨.

[6] GARETH J., DANIELA W., TREVOR H., ROBERT T., -**An Introduction to Statistical Learning with Applications in R.** New York Heidelberg Dordrecht London, 426.

[7] خانطوماني آية، ٢٠١٧–**مقارنة التصـــنيف المتعدد باســـتخدام طريقة المركبات الأساسية (PCA) وطريقة تحليل التمايز الخطي لفيشر (Fisher-LDA).** جامعة حلب، سوريا، ١٤٩.

[8] YIZENG L., QING-SONG X., HONG-DONG L., DONG-SHENG C., 2011- **Support Victor Machines and Their Application in Chemistry and Biotechnology.** United State of America, 194.

[9]− Michael Negnevitsky, 2005- **Artificial Intelligence A Guide to Intelligent Systems**- 2ed **Edition.Addison-Wesley p435 .**

[10] Zurada M. Jacek (1995) , **"Introduction to Artificial Neural Systems"** , PWS.

[11]−Ben Coppin,2004-**Artificial Intelligence Illuminated-** Jones And Bartlett Publishers

[12] **شلاف** فاطمة،2019 _ **معالجة المعطيات المفقودة والشاذة من خلال تطوير خوارزمية شبكة ذكاء صنعي متقدمة_ جامعة حلب، سوريا، رسالة دكتوراه.**

[13] Sivanandam S N, Sumathi S, Deepa S N.,2006-**introduction to Neural Networks Using Matlab 6.0-**Tata Mcgraw_Hill Publishing Company Limited New Delhi.p549.

[14]M.Guthikonda S.,2005,**Kohonen Self-Organizing Maps**, Wittenberg University.

[15]Mehotra K & Mohan K & Ranka S.,1997, **LEMENTS OF ARTIFICIAL NEURAL NETWORKS. MIT Press .**

[16]J. Holmström **.,** 2002-**Growing Neural Gas Experiments with GNG, GNG with Utility and Supervised GNG.** Uppsala University Department of Information Technology Computer Systems