

Fault Classification in Electrical Power Transmission System Using Machine Learning Algorithms

Dr. George Isber*
Dr. Ali Alsayed**
Rama Alkhtiar***

(Received 29 / 2 / 2024. Accepted 21 / 5 / 2024)

□ ABSTRACT □

As high-voltage lines are an important component of the electrical power system, it is necessary to accurately classify faults on them, thus improving the quality of electrical power and increasing the stability of the system.

In this research, machine learning algorithms were relied upon to classify faults in electrical power transmission lines due to their reliability and accuracy of their results. The performance of two machine learning algorithms was compared: Decision Tree (DT) and Support Vector Machine (SVM) in distinguishing fault conditions. From the normal working condition and determining the type of fault on a 400 kV high tension line. The line was modeled and faults simulated in the SIMULINK environment in MATLAB, to generate fault type data at different values of fault resistance. The data was filtered and processed in a Python environment to be used in training these algorithms.

The results showed that the DT algorithm outperformed the SVM algorithm in fault classification under different system conditions..

Keywords: electrical power transmission system, high voltage lines, type of fault, machine learning

Copyright



:Tishreen University journal-Syria, The authors retain the copyright under a CC BY-NC-SA 04

* Professor - Department of Electrical Power - Faculty of Mechanical and Electrical Engineering - Tishreen University - Lattakia – Syria. George.Isber@yahoo.com

**Associate Professor - Department of Electrical Power - Faculty of Mechanical and Electrical Engineering - Damascus University - Damascus – Syria. dr.alsayed@yahoo.com

*** Master Student - Department of Electrical Power - Faculty of Mechanical and Electrical Engineering - Tishreen University - Lattakia – Syria. rama.alkhtiar@tishreen.edu.sy

تصنيف العطل في نظام نقل الطاقة الكهربائية باستخدام خوارزميات تعلم الآلة

د. جورج اسبر*

د. علي السيد**

راما الختبار***

(تاريخ الإيداع 2024 / 2 / 29. قَبْلُ للنشر في 2024 / 5 / 21)

□ ملخص □

باعتبار خطوط التوتر العالي مكون هام في نظام القدرة الكهربائية، فإنه من الضروري تصنيف العطل عليها بدقة، وبالتالي تحسين جودة الطاقة الكهربائية، وزيادة استقرار النظام. تم في هذا البحث الاعتماد على خوارزميات تعلم الآلة من أجل تصنيف عطل خطوط نقل الطاقة الكهربائية نظراً لموثوقيتها ودقة نتائجها، حيث تم مقارنة أداء خوارزميتين لتعلم الآلة وهي: شجرة القرار (DT) و Decision Tree وآلة متجه الدعم (SVM) Support Vector Machine في تمييز حالة العطل من حالة العمل الطبيعية وتحديد نوع العطل على خط توتر عالي 400 kV. تمت نمذجة الخط ومحاكاة الأعطال في بيئة SIMULINK في برنامج MATLAB، وذلك لتوليد بيانات نوع العطل عند قيم مختلفة لمقاومة العطل. تم ترشيح البيانات ومعالجتها في بيئة Python ليتم استخدامها في تدريب هذه الخوارزميات. بينت النتائج تفوق أداء خوارزمية DT على خوارزمية SVM في تصنيف العطل وذلك عند شروط النظام المختلفة.

الكلمات المفتاحية: نظام نقل الطاقة الكهربائية، خطوط التوتر العالي، نوع العطل، خوارزميات تعلم الآلة



حقوق النشر : مجلة جامعة تشرين- سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص

CC BY-NC-SA 04

*أستاذ- قسم الطاقة الكهربائية- كلية الهندسة الميكانيكية والكهربائية- جامعة تشرين- اللاذقية- سورية.

George.Isber@yahoo.com

**أستاذ مساعد- قسم الطاقة الكهربائية- كلية الهندسة الميكانيكية والكهربائية- جامعة دمشق-دمشق- سورية

dr.alsayed@yahoo.com

***طالبة ماجستير- قسم الطاقة الكهربائية- كلية الهندسة الميكانيكية والكهربائية- جامعة تشرين- اللاذقية- سورية

rama.alkhtiar@tishreen.edu.sy

مقدمة:

تعد خطوط نقل الطاقة الكهربائية مكون أساسي من مكونات نظام القدرة الكهربائية، وهي تمتد لمسافات طويلة مما يجعلها عرضة لتأثير العديد من العوامل الخارجية مثل اتساخ العازل، والبرق، والرياح وما إلى ذلك، والتي تؤدي إلى حدوث أعطال. إذ تشكل أعطال خطوط النقل 80-90% من أعطال النظام، لذلك من المهم جداً تحديد نوع وموقع العطل لحماية خطوط النقل. يعد تصنيف أعطال القصر في خط النقل مهماً جداً للكشف الصحيح عن العطل وبالتالي يتم تقليل الأضرار التي تحدث في الأنظمة الكهربائية، وبالنتيجة زيادة موثوقية نظام القدرة الكهربائية [1].

تستخدم الطرق الآتية لتصنيف الأعطال في خطوط النقل الكهربائية [1]:

- ❖ الطريقة المعتمدة على قياس الممانعة.
- ❖ الطريقة المعتمدة على الأمواج المسافرة.
- ❖ طرق الذكاء الاصطناعي.

تعتمد طريقة قياس الممانعة بشكل أساسي على مطالات الجهود والتيارات عند التردد الأساسي، تعد هذه الطريقة بسيطة و رخيصة، ولكن أداءها يتأثر بمقاومة العطل وبارامترات المصدر [2]

إن الطريقة المعتمدة على الموجة المسافرة تتطلب تردد تقطيع أعلى من 1MHz والذي يصعب تنفيذه في المجال العملي كما لديها مشاكل للتمييز بين الأمواج المنعكسة من العطل ومن الطرف الآخر للخط [2, 3].

ذكر الباحثون في المرجع [4] أن الطرق التقليدية تعطي نتائج غير دقيقة في تصنيف العطل وتحديد موقعه على خطوط نقل الطاقة الكهربائية، فقد تؤدي إلى فصل الجزء السليم من الخط، لذلك توجهوا حديثاً إلى استخدام طرق الذكاء الاصطناعي وخوارزميات تعلم الآلة التي تتميز بدقتها العالية.

صنف الباحثون في المرجع [5] نوع العطل وحددوا موقعه على خط توتر 380 kv بطول 360 km مغذى من جهة واحدة باستخدام خوارزمية SVM وتحويل الموجة المتقطع، وقد توصلوا إلى خطأ تصنيف أقل من 1% لجميع أنواع الأعطال المختبرة، كما ذكر الباحثون في المرجع [6] أن استخدام خوارزمية SVM يُحسن أداء الحماية التقليدية لخطوط النقل. أشار الباحث في المرجع [7] أن خوارزمية SVM يمكنها تحديد الأصناف بدقة مقارنة بالنظم الخبيرة والشبكة العصبية الاصطناعية ANN والمنطق الضبابي.

حدد الباحثون في المرجع [8] نوع العطل على خط نقل مزدوج الدارة باستخدام خوارزمية شجرة القرار، إذ استخدموا إشارات جهود وتيارات الأطوار المقاسة من طرف واحد كمدخل لها، أظهرت النتائج قدرة الخوارزمية على تصنيف العطل بدقة عالية في أقل من ربع دورة. قارن الباحثون في المرجع [9] أداء خوارزمية شجرة القرار مع شبكة عصبونية اصطناعية ANN مدربة بطريقة الانتشار الخلفي للخطأ Back Propagation في تصنيف العطل وتحديد موقعه على خط توتر 400 kv مغذى من جهة واحدة فقط. أثبتت النتائج تفوق أداء خوارزمية شجرة القرار على الشبكة العصبونية المقترحة بدقة وصلت إلى 99.97 و 99.88% لكل منهما على التوالي. حدد الباحثون في المرجع [10] نوع العطل وموقعه على خط توتر 220 kv مغذى من طرف واحد باستخدام خوارزمية الغابة العشوائية ولكن لم يؤخذ بالاعتبار التمييز بين حالة العطل وحالة العمل الطبيعية. في المرجع [11] حددت الباحثة نوع وموقع العطل على خط نقل طاقة كهربائية مغذى من طرفيه باستخدام خوارزميات تعلم الآلة، ولكن لم يؤخذ أيضاً بالاعتبار التمييز بين حالة العمل الطبيعية وحالة العطل.

تم في هذا البحث مقارنة أداء خوارزميات DT و SVM في تحديد نوع العطل على خط توتر 400 kv بطول

300 Km مغذى من طرفين، وذلك عند أنواع أعطال مختلفة: (طور مع الأرض، ثنائية الأطوار دون الأرض، ثنائية الأطوار مع الأرض، ثلاثية الأطوار، حالة العمل الطبيعية). وذلك باستخدام بيئة "Python" البرمجية.

أهمية البحث وأهدافه:

تأتي أهمية البحث من تطبيق أحد تقنيات الذكاء الاصطناعي (خوارزميات تعلم الآلة) في حل أحد مشاكل نظم القدرة الكهربائية (تصنيف العطل بدقة عالية على خطوط نقل الطاقة الكهربائية) وذلك بهدف رفع موثوقية نظام القدرة الكهربائية.

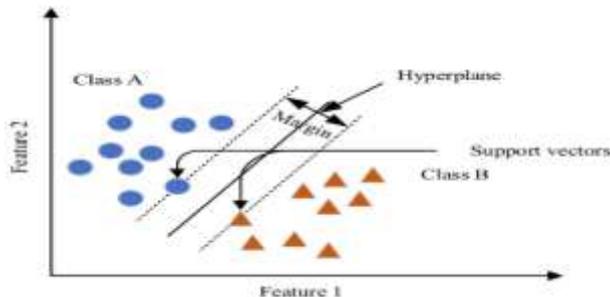
طرائق البحث ومواده:

تم في هذا البحث نمذجة ومحاكاة نظام نقل طاقة كهربائية ثلاثي الأطوار 400 kV في برنامج MATLAB، وتم استخدام خوارزميات SVM وشجرة القرار لتحديد نوع العطل على هذا الخط، حيث دخل الخوارزمية هو مطالات جهود وتيارات الأطوار عند التردد الأساسي لشبكة التغذية 50 Hz وذلك عند حالة العمل الطبيعية وعند حدوث عطل على الخط ببارامترات مختلفة (مقاومة العطل ومسافته)، أما خرج الخوارزمية هو التمييز بين حالة العطل وحالة العمل الطبيعية مع تحديد نوع العطل.

خوارزمية آلة متجه الدعم (Support Vector Machine (SVM))

تتتمي خوارزمية آلة المتجهات الداعمة إلى خوارزميات تعلم الآلة الخاضعة للإشراف. تستخدم لأجل مهام التصنيف والانحدار. تعمل هذه الخوارزمية على تصنيف البيانات القابلة للفصل خطياً من خلال إيجاد خط فصل يدعى hyper plane بحيث يكون فيها ما يعرف بهامش الفصل margin بين عينات التصنيف المختلفة أكبر ما يمكن. يشير مصطلح أشعة متجهات الدعم إلى نقاط البيانات التي تتوضع على الحدود الفاصلة بين أصناف البيانات والتي تستخدم لدعم الموقع المثالي لمستوي الفصل [12].

يبين الشكل (1) مبدأ مصنف SVM عند فصل نوعين من البيانات القابلة للفصل خطياً:

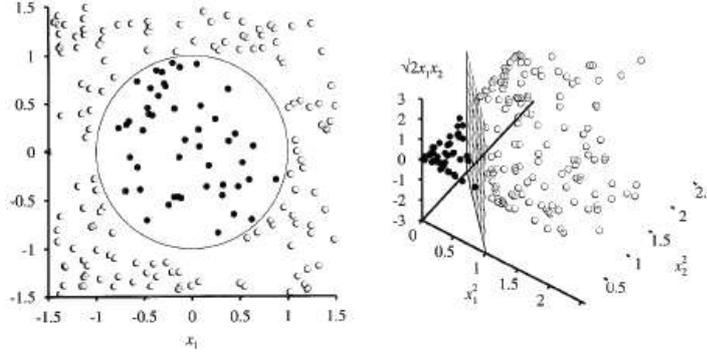


الشكل (1) مصنف SVM [12]

عند زيادة عدد الأصناف تزداد مشكلة الخطية في مصنف SVM، حيث يعبر عن ذلك ببيانيا من خلال وجود نقاط تقاطع مشتركة بين بيانات الأصناف [13]. في هذه الحالة لا يوجد إمكانية لفصل البيانات باستخدام خط فصل في فضاء ثنائي البعد عند ذلك يتم تعيين متجهات الميزة من خلال نقل البيانات غير الخطية إلى فضاء أبعاد أكبر كما في الشكل (2) تكون فيه قابلة للفصل خطياً. نطلق على التابع المسؤول عن تحقيق هذه العملية اسم تابع النواة [14] في بحثنا استخدمنا أحد أهم وأكثر توابع النواة استخداماً Gaussian Radial Basis Function والذي يعطى بالمعادلة [14]:

$$K(X, X') = \exp\left(-\frac{\|X-X'\|^2}{2\sigma}\right) \dots\dots\dots (1)$$

حيث σ : الانحراف المعياري لتوزيع غوص.



الشكل (2) فصل العينات في قضاء ثلاثي الأبعاد [13]

من مزايا استخدام خوارزمية SVM [15]:

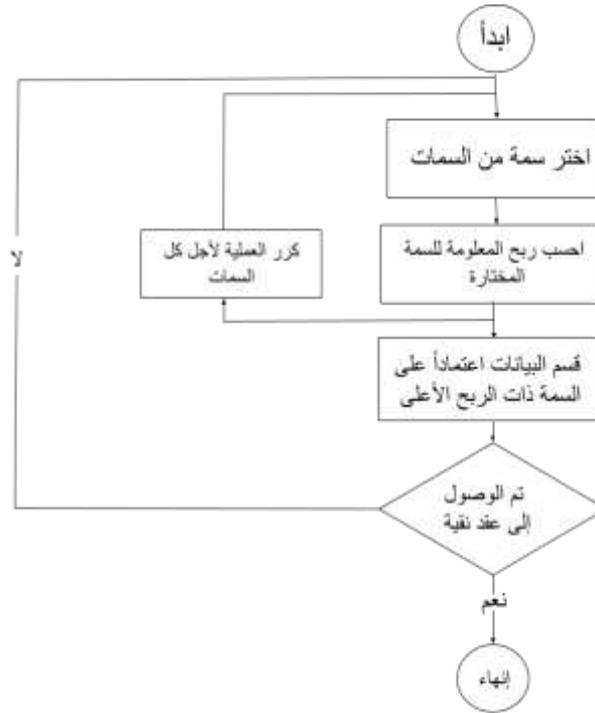
1- قابلة للتعميم بشكل جيد، بمعنى بمجرد تدريبها على مجموعة البيانات ومن ثم إعطاؤها بيانات اختبار جديدة قادرة على التصنيف الصحيح بدرجة كبيرة.

خوارزمية شجرة القرار (Decision Tree (DT))

تعتبر شجرة القرار جزء من عائلة خوارزميات تعلم الآلة الخاضعة للإشراف يمكن استخدامها لحل مشاكل التصنيف والانحدار. هدفها الأساسي هو بناء نموذج تدريب يمكن استخدامه للتنبؤ بصنف أو قيمة المتغير الهدف من خلال تعلم قواعد القرار المستنبطة من بيانات التدريب.

تعتمد هذه التقنية على نموذج الشجرة حيث تتكون من عقد وفروع وأوراق، يتم تقييم فعالية كل سمة دخل في عملية بناء الشجرة في تحديد صنف الخرج عن طريق معيار محدد مسبقاً (ربح المعلومات)، ليتم عندها اختيار أكثر السمات فعالية (التي تحقق أعلى ربح معلومات) كعقدة بداية والتي تتفرع إلى قسمين بناءً على قيمة حدية ما، ثم لأجل كل فرع يتم إيجاد السمة التالية الأكثر فعالية وتكرر عملية التفرع حتى الوصول إلى عقدة نهائية تحتوي على صنف خرج واحد [16].

يوضح الشكل (3) مبدأ عمل شجرة القرار:



الشكل (3) مبدأ عمل شجرة القرار

1. ربح المعلومات والانتروبي (Entropy and Information Gain):

الانتروبي: هو مقياس للعشوائية في مجموعة البيانات. يأخذ القيم بين 0 و 1 ، حيث كلما اقتربت قيمته من الصفر كلما كان أفضل [16].
يعطى الانتروبي بالعلاقة [16]:

$$\text{Entropy}(S) = -\sum_{i=1}^c p_i \log 2^{p_i} \dots\dots\dots(2)$$

حيث:

S: مجموعة بيانات التدريب.

C: عدد الأصناف وهو في بحثنا 11 صنف.

p: هو احتمال الاختيار العشوائي لعينة في الصنف i

ربح المعلومات: هو أحد المقاييس المستخدمة للتجزئة ويسمى غالباً بالمعلومات المتبادلة. وهو على عكس الانتروبي كلما زادت قيمته كلما كان أفضل.

يعطى الربح المعتمد على مفهوم الانتروبي بالعلاقة التالية [17]:

$$\text{ربح المعلومة} = \text{الانتروبي للأب} - \text{المجموع الموزون لأنتروبي الأبناء} \dots\dots\dots(3)$$

من مزايا خوارزمية شجرة القرار:
سهولة الفهم والتفسير.

مراحل العمل:

جرت عملية المعالجة وفق المراحل الآتية:

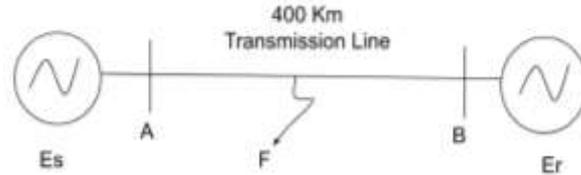
1- بناء نموذج خط نقل ثلاثي الطور بطول 300 Km، مغذى من منبعين بتوتر 400 kV، وبتردد 50 Hz في بيئة MATLAB –SIMULINK.

تعطى بارامترات النظام المدروس بالجدول (1) [2]:

الجدول (1) بارامترات نظام نقل الطاقة المدروس

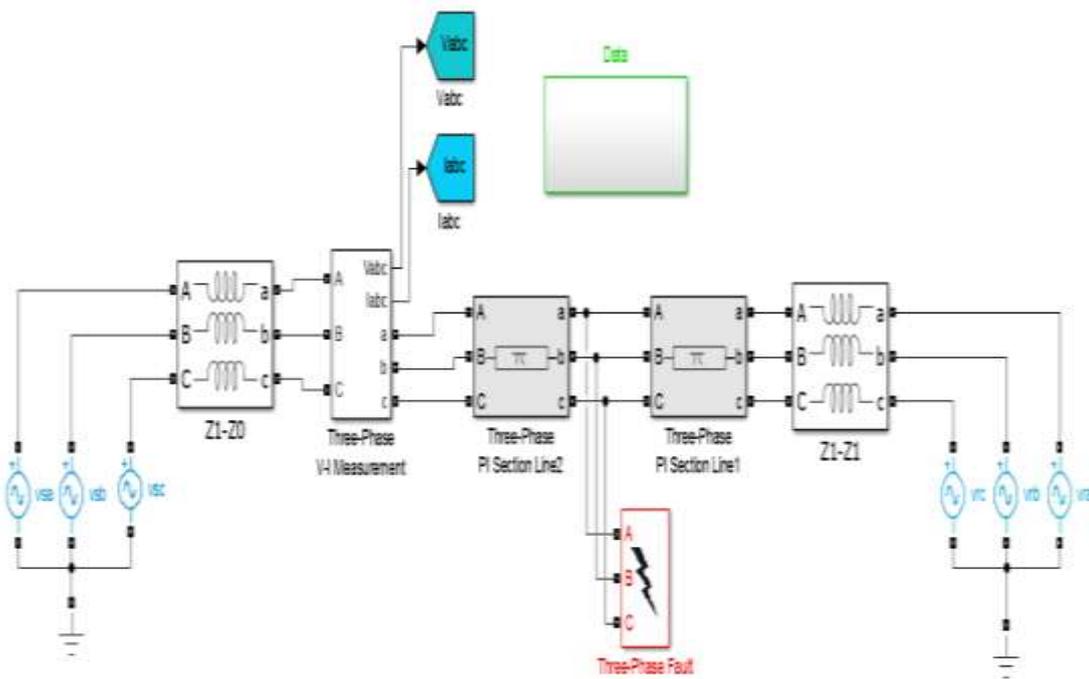
| بارامترات منبع التغذية | |
|-----------------------------------|------------------------------|
| ممانعة التتابع الموجب Z_1 | $1.31+j16$ [Ω] |
| ممانعة التتابع الصفري Z_0 | $2.22+j27.6$ [Ω] |
| بارامترات خط التوتر الثلاثي الطور | |
| ممانعة التتابع الموجب Z_1 | $8.15+j94.5$ [Ω /Km] |
| ممانعة التتابع الصفري Z_0 | $92.5+j308$ [Ω /Km] |
| سعة التتابع الموجب C_1 | 14 [nF/Km] |
| سعة التتابع الصفري C_0 | 7.5 [nF/Km] |

يبين الشكل (4) مخطط الخط الواحد (Single Line Diagram) لنظام نقل القدرة الكهربائية المدروس:



الشكل (4) مخطط الخط الواحد (Single Line Diagram) لنظام نقل القدرة المدروس

يبين الشكل (5) نموذج النظام المدروس في بيئة MATLAB–SIMULINK:



الشكل (5) نموذج النظام المدروس في بيئة MATLAB-SIMULINK

2- محاكاة الأعطال وإعداد بيانات التدريب والاختبار:

جرت محاكاة الأعطال التالية:

AG-BG-CG-AB-BC-AC-ABG-BCG-ACG-ABC

حيث:

A: الطور الأول.

B: الطور الثاني.

C: الطور الثالث.

G: الأرض.

تم تجهيز بيانات التدريب والاختبار بتغيير كل من مسافة العطل (L) ومقاومته (R_f) عند كل نوع من أنواع الأعطال السابقة الذكر كما هو مبين في الجدول (2) حيث بلغ حجم قاعدة البيانات الكلية 74371 عينة (71971 عينة عطل و 2400 حالة عمل طبيعية).

الجدول (2) بيانات التدريب والاختبار

| | |
|---|---------------------------------|
| AG-BG-CG-AB-BC-AC-ABG-BCG-ACG-ABC- No fault | نوع العطل N |
| 0.01, 10, 30 | مقاومة العطل R_f [Ω] |
| 0.125:300 Step: 0.125 | مكان حدوث العطل L [Km] |

- 3- قياس إشارات تيارات وجهود الأطوار عند قضيبي التجميع A عند حدوث العطل، ثم ترشيحها بمرشح الترددات الزائفة (Anti-Aliasing Filter (AAF) وتقطيعها بتردد تقطيع (Sampling Frequency) مساوٍ لـ 800 Hz .
- 4- الحصول على مطال التردد الأساسي (fundamental frequency phasor) لإشارات جهود وتيارات الأطوار عند حدوث الأعطال لكل حالة من الحالات المختلفة السابقة باستخدام تحويل فورييه المنقطع . Discrete Fourier Transform (DFT)
- 5- تخزين بيانات العطل بصيغة EXCEL حيث يشكل كل سطر من مجموعة البيانات عينة تتألف من مجموعة سمات وهي على التوالي: مطالات الجهود (V1-V2-V3) والتيارات (I1-I2-I3) عند التردد الأساسي 50 Hz المأخوذة عند حالة العطل وحالة العمل الطبيعية كما يبين الجدول (3).
تستخدم هذه القيم في تدريب واختبار الخوارزميات المقترحة حيث تمثل مطالات الجهود والتيارات دخل الخوارزميات ويمثل التمييز بين حالة العمل الطبيعية وحالة العطل خرجها.

الجدول (3) بنية بيانات التدريب والاختبار

| V1 | V2 | V3 | I1 | I2 | I3 | Class |
|------------|------------|------------|---------|---------|---------|----------|
| 274872 | 320082.1 | 365562.1 | 8720.53 | 546.53 | 793.354 | AG |
| 252803.84 | 262207.15 | 339961.77 | 4394.57 | 4760.7 | 711.54 | ABG |
| 268703.95 | 321577.33 | 363594.4 | 8301.56 | 556.09 | 791 | AG |
| 329204.760 | 329207.616 | 329206.266 | 709.369 | 709.376 | 709.372 | NO fault |
| 329861.18 | 199294.27 | 199908.69 | 710.061 | 10203 | 10898 | BC |
| 329204.76 | 329207.61 | 329206.266 | 709.369 | 709.376 | 709.372 | NO fault |

6- تقسيم البيانات إلى 70% من أجل تدريب الخوارزميات و30% من أجل اختبارها.

7- تنفيذ الخوارزميات باستخدام لغة البرمجة Python ضمن بيئة Google Colab.

معايير تقييم الأداء:

تم في هذا البحث تقييم أداء خوارزميات تعلم الآلة المستخدمة من أجل تصنيف العطل على خط نقل الطاقة الكهربائية المدروس باستخدام مصفوفة الارتباك (Confusion Matrix)، التي تعد من أفضل المقاييس في مجال التحليل التنبؤي [13]، حيث يمكننا من خلال هذه المصفوفة معرفة دقة النموذج (accuracy) التي تعبر عن صحة أدائه وتعرف بأنها نسبة عدد العينات التي تم توقعها بشكل صحيح على إجمالي عدد العينات، كما يمكننا من خلال هذه المصفوفة معرفة دقة تصنيف كل نوع على حدا، قيمة الخطأ المقابلة بالإضافة إلى عدد العينات التي تم تصنيفها بشكل صحيح أو خاطئ مع تحديد الصنف الذي تنتمي إليه العينات.

استخدمنا بالإضافة إلى المعيار السابق بارامتر إضافي وهو زمن تحديد نوع العطل والذي يعبر عن سرعة الحصول على خرج الخوارزمية المدربة لتقييم أداء الخوارزميات المستخدمة في هذا البحث.

النتائج والمناقشة:

جرى في هذا البحث مقارنة أداء خوارزمتين من خوارزميات تعلم الآلة وهما خوارزمية آلة متجهات الدعم وخوارزمية شجرة القرار في تمييز حالة العطل من حالة العمل الطبيعية وتحديد نوع العطل على خط توتر عالي 400 kV مغذى من طرفيه، وتم التوصل إلى الآتي:

تحديد نوع العطل باستخدام خوارزمية SVM:

من أجل تنفيذ خوارزمية SVM تم استخدام تابع RBF كنواة يمكنها نقل فضاء السمات السداسي (مطالات الجهود و التيارات) إلى فضاء عالي الأبعاد 'RBF' Kernel ، كذلك حصلنا على أفضل النتائج عند ضبط بارامتر عامل التنظيم المتحكم بعرض هامش الفصل عند القيمة $C = 20$ ، وعند قيمة $\sigma = 1/10$.
يبين الشكل (6) مصفوفة الارتباك لخوارزمية SVM في تحديد نوع العطل التي حصلنا عليها.

| | AG | BG | CG | AB | ABG | BC | BCG | AC | ACG | ABC | No fault | sum_col |
|----------|---------------|---------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------|--------------|-----------------|
| AG | 2214 9.92% | 0 0.0% | 0 0.0% | 44 0.20% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2258 98.05% |
| BG | 0 0.0% | 2165 9.70% | 0 0.0% | 0 0.0% | 0 0.0% | 39 0.17% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2204 98.23% |
| CG | 0 0.0% | 0 0.0% | 2113 9.47% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 54 0.29% | 0 0.0% | 0 0.0% | 0 0.0% | 2177 97.06% |
| AB | 0 0.0% | 0 0.0% | 0 0.0% | 2101 9.42% | 108 0.48% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2209 95.11% |
| ABG | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 1976 8.86% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 1976 100% |
| BC | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2103 9.43% | 107 0.48% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2210 95.16% |
| BCG | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 1987 8.91% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 1987 100% |
| AC | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2189 9.81% | 111 0.50% | 0 0.0% | 0 0.0% | 2300 95.17% |
| ACG | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2067 9.26% | 0 0.0% | 0 0.0% | 2067 100% |
| ABC | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2189 9.81% | 0 0.0% | 2189 100% |
| No fault | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 734 3.29% | 734 100% |
| sum_col | 2214 100% | 2165 100% | 2113 100% | 2145 97.95% | 2084 94.82% | 2142 98.18% | 2094 94.89% | 2253 97.16% | 2178 94.90% | 2189 100% | 734 100% | 22311 97.88% |
| | 0.00% | 0.00% | 0.00% | 2.05% | 5.18% | 1.82% | 5.11% | 2.84% | 5.10% | 0.00% | 0.00% | 2.12% |
| Actual | AG | BG | CG | AB | ABG | BC | BCG | AC | ACG | ABC | No fault | sum_lin |

الشكل (6) مصفوفة الارتباك لخوارزمية SVM في تحديد نوع العطل

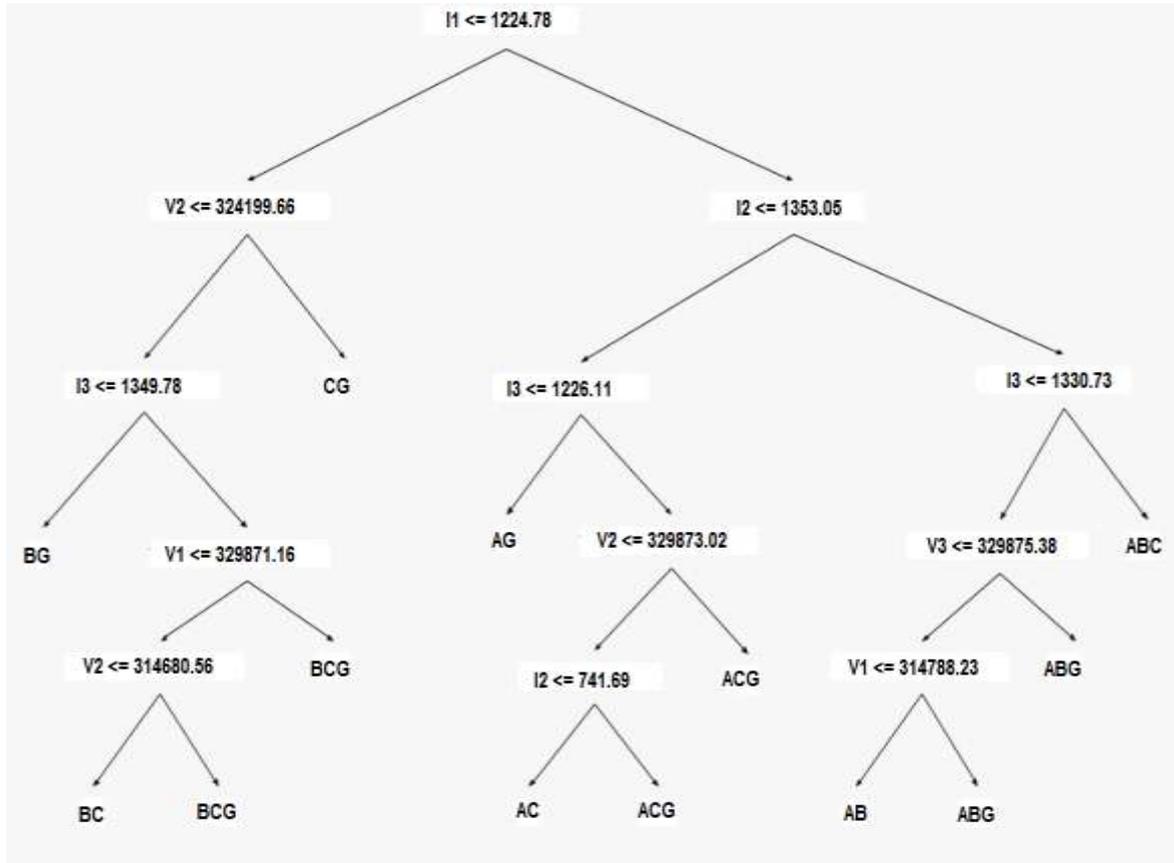
نلاحظ من الشكل (6) أن الدقة العامة للنموذج وصلت إلى أكثر من 97% مع دقة تصنيف مثالية بلغت 100% لأجل الحالات AG, BG, CG, ABC, No fault أي أن جميع عينات هذه الأنواع تم توقعها بشكل صحيح. لأجل النوع AB بلغت دقة التصنيف 97.95%، إذ تم توقع 44 عينة بشكل خاطئ على أنها تنتمي للنوع AG وهو ما

يشكل نسبة 0.20% من مجموع بيانات الاختبار الكلية، بينما لأجل النوع ABG حصلنا على دقة تصنيف مقدارها 94.82% مع توقع خاطئ مقداره 108 عينة لأجل النوع AB بنسبة 0.48% من عدد عينات الاختبار، وفي النوع BC كانت دقة التصنيف 98.18%، إذ تم توقع 39 عينة بشكل خاطئ على أنها منتمية للصنف BG وهذا يشكل نسبة 0.17% من عينات الاختبار، ولأجل النوع BCG حصلنا على دقة تصنيف 94.89%، إذ بلغ عدد العينات الخاطئة 107 على أنها منتمية للنوع BC والتي تشكل 0.48% من عدد عينات الاختبار. في العطل من النوع AC بلغت دقة التصنيف 97.16% مع توقع 64 عينة بشكل خاطئ على أنها منتمية للصنف CG وهو ما يشكل 0.29% من عينات الاختبار بينما لأجل النوع ACG كانت دقة التصنيف 94.90% مع توقع خاطئ مقداره 111 عينة لأجل النوع AC وبالتالي تكون نسبة العينات الخاطئة 0.50% من العينات المستخدمة في عملية الاختبار، لتصبح بذلك نسبة الخطأ الكلية في نموذج SVM 2.12% فقط.

تحديد نوع العطل باستخدام خوارزمية DT:

لتنفيذ خوارزمية شجرة القرار قمنا باختيار الأنتروبي كطريقة لقياس جودة تقسيم الشجرة 'entropy = criterion'، كذلك تم استخدام القيمة الافتراضية لبارامتر عمق الشجرة والتي من خلالها يتم تقسيم الشجرة حتى الوصول إلى عقد نقية بالكامل تحتوي على صنف واحد فقط 'default = max_depth' .

يبين الشكل (7) شجرة التصنيف التي حصلنا عليها.



الشكل (7) شجرة التصنيف

كما هو واضح من الشكل (7) إن أفضل سمة في تقسيم البيانات وتشكيل الشجرة كانت مطال التيار الأول عند القيمة 1224.78 ثم لأجل كل فرع من الشجرة تستمر عملية اختبار أفضل السمات حتى الوصول إلى عقد نقية (تحوي صنف عطل واحد).

يبين الشكل (8) مصفوفة الارتباك لخوارزمية DT التي حصلنا عليها.

Confusion matrix

| | | | | | | | | | | | | |
|----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|------------------------|
| | AG | BG | CG | AB | ABG | BC | BCG | AC | ACG | ABC | No fault | sum_col |
| AG | 2214 9.92% | 0 0.0% | 0 0.0% | 2214 100% 0.00% |
| BG | 0 0.0% | 2165 9.70% | 0 0.0% | 0 0.0% | 2165 100% 0.00% |
| CG | 0 0.0% | 0 0.0% | 2113 9.47% | 0 0.0% | 0 0.0% | 2113 100% 0.00% |
| AB | 0 0.0% | 0 0.0% | 0 0.0% | 2145 9.61% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2145 100% 0.00% |
| ABG | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2084 9.34% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2084 100% 0.00% |
| BC | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2142 9.60% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2142 100% 0.00% |
| BCG | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2094 9.39% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 2094 100% 0.00% |
| AC | 0 0.0% | 2253 10.10% | 0 0.0% | 0 0.0% | 0 0.0% | 2253 100% 0.00% |
| ACG | 0 0.0% | 2178 9.76% | 0 0.0% | 0 0.0% | 2178 100% 0.00% |
| ABC | 0 0.0% | 2189 9.81% | 0 0.0% | 2189 100% 0.00% |
| No fault | 0 0.0% | 734 3.29% | 734 100% 0.00% |
| sum_col | 2214 100% 0.00% | 2165 100% 0.00% | 2113 100% 0.00% | 2145 100% 0.00% | 2084 100% 0.00% | 2142 100% 0.00% | 2094 100% 0.00% | 2253 100% 0.00% | 2178 100% 0.00% | 2189 100% 0.00% | 734 100% 0.00% | 22311 100% 0.00% |
| | AG | BG | CG | AB | ABG | BC | BCG | AC | ACG | ABC | No fault | sum_lin |

الشكل (8) مصفوفة الارتباك لخوارزمية شجرة القرار

نلاحظ من الشكل (8) أننا حصلنا على دقة عامة للنموذج مقدارها 100%، وبالتالي دقة مقدارها 100% لأجل جميع أنواع الأعطال، الأمر الذي يعني أن جميع العينات تم توقعها بشكل صحيح دون خطأ. ويبين الجدول (4) الزمن الذي استغرقته كل خوارزمية في تصنيف العطل.

الجدول (4) مقارنة أداء الخوارزميات من حيث زمن تحديد نوع العطل

| | |
|---------------------------|-------------|
| زمن تحديد نوع العطل [sec] | الخوارزمية |
| 0.00169 | SVM |
| 0.000813 | شجرة القرار |

نلاحظ من الجدول (4) أن:

- أزمنة تحديد نوع العطل صغيرة جداً إذ أعطت شجرة القرار أقل زمن تصنيف وهو 0.000813 ثانية.

الاستنتاجات والتوصيات:**الاستنتاجات:**

أظهرت الدراسة قدرة خوارزميات تعلم الآلة المدروسة على تحديد نوع العطل بدقة عالية عند شروط النظام المختلفة مما يزيد من موثوقية نظام القدرة الكهربائية، حيث:

✓ أعطت الخوارزميات المستخدمة دقة تحديد عالية لنوع العطل تراوحت بين 97.77% و 100 % وبالتالي يمكن استخدام هذه الخوارزميات في عملية تصنيف نوع العطل على خطوط نقل الطاقة الكهربائية باعتبارها خوارزميات ذات دقة تصنيف عالية.

✓ أزمنة تحديد نوع العطل للخوارزميات صغيرة جداً بالتالي هذه الخوارزميات مناسبة للعمل في الزمن الحقيقي.
 ✓ بمكاملة عامل الدقة مع زمن التصنيف نلاحظ تفوق أداء خوارزمية شجرة القرار في تحديد نوع العطل بدقة عالية عند شروط النظام المختلفة مقارنة بخوارزمية SVM، حيث:
 ✓ حققت خوارزمية شجرة القرار الأداء الأفضل (أعلى دقة 100% وأسرع زمن تصنيف مقداره 0.000813 ثانية).

التوصيات:**نوصي بتطوير هذا البحث من خلال:**

❖ بناء واجهة برمجية لنموذج شجرة القرار بحيث تسهل على العاملين استخدام هذا النموذج في محطات الطاقة الكهربائية من أجل تصنيف عطل خطوط نقل الطاقة الكهربائية.

References:

1. Tuba, A. Using Machine Learning Algorithms For Classifying Transmission Line Faults. Dicle Universitesi Mühendislik Fakültesi Mühendislik Dergisi. 2022;13(2):227–34.
2. Ray, P. Mishra, DP. Support vector machine based fault classification and location of a long transmission line. 2016;
3. Ray, P. Panigrahi, BK. Senroy. Hybrid methodology for fault distance estimation in series compensated transmission line. 2013. IET_Generation, Transmission & Distribution.;7(5):431–9
4. Mehdi, A. Hussain, A. Haider, W.Hassan, S. M. Saad, Chul-Hwan Kim. Classification of power system faults using Random Forest classifier. 2020. Conference Paper.
5. Ekici, S. Support Vector Machines for classification and locating faults on transmission Lines.2012. Elsevier. pp. 1650–1658
6. V, Malathi. N,S,Marimuthu. Multi-class Support Vector Machine Approach for Fault classification in Power Transmission Line.2008. IEEE, PP. 67-71
7. Lilik, J. Awalın, Kanendra, N. Hadi, S. Fault Types Classification Using Support Vector Machine (SVM).2019. AIP Conference Proceedings 2129, 020132
8. Jamehbozorg, A. Shahrtash, M. A Decision Tree-Based Method for Fault Classification in Double-Circuit Transmission Lines.2010. IEEE Transactions on Power Delivery, 25(4):2184 – 218
9. Upendar, J. Gupta, CP. Singh, GK. Statistical decision-tree based fault classification scheme for protection of power transmission lines. International Journal of Electrical Power & Energy Systems.2012. Electrical Power and Energy Systems;36(1):1-12
10. Alkhtiar, R. Isber, G. Alsayed, A. Fault classification and location on high voltage lines using the random forests algorithm. Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (45) No. (1) 2023

11. Alkhtiar, R. Comparing the performance of the most important machine learning algorithms in determining the type and location of faults on high voltage lines [Master]. [Syria]: Tishreen University.2023
12. Yang, H. Liu, X. Zhang, D. Chen, T. Li, C. Huang W. Machine learning for power system protection and control.2021. The Electricity Journal 34(1):106881
13. Da'ada, J. Improvement and classification of epileptic seizures from brain signaling [Master]. [Syria]: Tishreen University.2021
14. Somvanshi, M., Chavan, P. Tambade,S. A review of machine learning techniques using decision tree and support vector machine, 2016 International Conference on Computing Communication Control and automation (ICCUBEA), 1-7.
15. Samantaray, S.R., Dash, P.K., Panda, G. Power System Protection Using Machine Learning Technique.2010. Computational Intelligence in Power Engineering pp 171–198
16. Jijo, B. Abdulazeez, A. Classification Based on Decision Tree Algorithm for Machine Learning.2021. JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS. Vol. 02, No. 01, pp. 20 – 28
17. Venkatesh, V. Fault Classification and Location Identification on Electrical Transmission Network Based on Machine Learning Methods.2018. Virginia Commonwealth University Richmond