|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **WYDZIAŁ ELEKTRYCZNY**Kierunek: elektrotechnika**Specjalność: Elektroenergetyka** **Rafał Zaręba**  Nr albumu: 238657  **Praktyczna implementacja sztucznych sieci neuronowych w celu diagnostyki aparatury przeciwprzepięciowej w czasie rzeczywistym**  *Practical implementation of artificial neural networks for the purpose of real-time surge arresters diagnostics*  **INŻYNIERSKA PRACA DYPLOMOWA**  Studia: niestacjonarne  *Opiekun:* **Dr inż. Paweł Kostyła**  Katedra <nazwa Katedry>   |  |  | | --- | --- | | ............... | .......................... | | *ocena* | *data, podpis opiekuna* | |
| **Wrocław, 2020** |

Spis treści

[Spis treści 2](#_Toc38258276)

[1. WSTĘP 3](#_Toc38258277)

[2. CEL I ZAKRES PRACY 4](#_Toc38258278)

[2.1. Cel pracy 4](#_Toc38258279)

[2.2. Architektura modelu 4](#_Toc38258280)

[2.3. Użyte narzędzia i technologie 4](#_Toc38258281)

[3. WSTĘP TEORETYCZNY 5](#_Toc38258282)

[3.1. Konserwacja predykcyjna 5](#_Toc38258283)

[3.2. Sztuczne sieci neuronowe 6](#_Toc38258284)

[3.2.1. Sieci neuronowe, a detekcja anomalii. 7](#_Toc38258285)

[3.3. Warystory tlenkowo-cynkowe 8](#_Toc38258286)

[3.3.1. Budowa 8](#_Toc38258287)

[3.3.2. Diagnostyka 9](#_Toc38258288)

[4. ZAPROPONOWANE ROZWIĄZANIE 11](#_Toc38258289)

[4.1. Rekurencyjne sieci neuronowe 11](#_Toc38258290)

[4.1.1. Budowa 12](#_Toc38258291)

[4.1.2. Wyjście sieci rekurencyjnej 13](#_Toc38258292)

[4.1.3. Problem zanikających gradientów. 14](#_Toc38258293)

[4.1.4. Sieci LSTM 16](#_Toc38258294)

[4.1.5. Sieci dwukierunkowe LSTM 18](#_Toc38258295)

[4.2. Autoenkodery 19](#_Toc38258296)

[4.3. Opis zaprojektowanego algorytmu 20](#_Toc38258297)

[5. PRAKTYCZNA IMPLEMENTACJA 22](#_Toc38258298)

[5.1. Wczytanie i przygotowanie danych uczących 22](#_Toc38258299)

[5.2. Definicja modelu sieci neuronowej 22](#_Toc38258300)

[5.3. Uczenie 22](#_Toc38258301)

[5.4. Walidacja 22](#_Toc38258302)

[6. WNIOSKI 22](#_Toc38258303)

# WSTĘP

W dobie powszechnego dostępu do zasobów komputerowych wiele gałęzi gospodarki uległo cyfryzacji. Wyjątkiem nie jest także cały sektor elektroenergetyczny. Niezawodność jego elementów, dobrze zaprojektowana automatyzacja oraz umiejętna diagnostyka działających w nim urządzeń jest kluczowa dla poprawnego funkcjonowania całego systemu. Wszechobecność wszelkiego rodzaju aparatury pomiarowej oraz czujników, nierzadko komunikujących się ze sobą powoduje, że w każdej chwili generowane są ogromne ilości danych. Rozwój technologii komputerowych umożliwił nie tylko ich łatwe pozyskanie, lecz także niemal nieograniczone i nieprzerwane gromadzenie. Szacuje się, że z upływem każdych dwóch lat sumaryczna ilość danych wygenerowana i zarejestrowana przez człowieka podwaja się. Podczas gdy dane same w sobie okazują się być bardzo wartościowe, istnieje duża potrzeba rozwijania nowoczesnych technik ich przetwarzania. Dopiero odpowiednio przygotowane i ustrukturyzowane dane mogą wspomóc wiele procesów technologicznych i biznesowych. Coraz częściej wykorzystuje się w tym celu algorytmy uczenia maszynowego i sztuczne sieci neuronowe.

Szczególnie interesującą i innowacyjną dziedziną uczenia maszynowego, która prężnie rozwija się na fali zainteresowania sztucznymi sieciami neuronowymi jest detekcja anomalii (ang. anomaly detection). Obejmuje między innymi takie zagadnienia jak wykrywanie podejrzanych transakcji płatniczych (ang. fraud detection), detekcję niebezpiecznych zdarzeń w ruchu internetowym, diagnozę chorób, czy wykrywanie nadchodzących nieprawidłowości w działaniu urządzeń lub maszyn, których parametry mierzone są w czasie rzeczywistym.

Zagadnieniem wykorzystania sztucznych sieci neuronowych w celu diagnostyki urządzeń w czasie rzeczywistym oraz przykładem praktycznej implementacji takich sieci zajęto się w niniejszej pracy. Jako badane urządzenie wykorzystano warystor tlenkowo-cynkowy spełniający rolę ogranicznika przepięć w systemie elektroenergetycznym.

# CEL I ZAKRES PRACY

## Cel pracy

Celem pracy jest zaprojektowanie i praktyczna implementacja modelu uczenia maszynowego opartego o architekturę sztucznych sieci neuronowych. Zadaniem modelu jest monitorowanie stanu technicznego warystora tlenkowo-cynkowego w czasie rzeczywistym na podstawie odczytanych próbek napięcia oraz natężenia prądu. Ponadto model powinien samoczynnie sygnalizować użytkownikowi o pogorszeniu się stanu warystora i z wyprzedzeniem podać informację o zbliżającej się degradacji jego struktury w celu dokonania jego wymiany.

## Architektura modelu

Model zbudowany został w oparciu o architekturę głębokiego autoenkodera z rekurencyjnymi warstwami LSTM. Detekcja występujących w urządzeniu anomalii oparta jest o błąd rekonstrukcji sekwencji danych wejściowych.

## Użyte narzędzia i technologie

Do realizacji projektu użyto wysokopoziomowego języka programowania Python w wersji 3.8. Do przygotowania i eksploracji danych użyto zoptymalizowanych bibliotek do obliczeń numerycznych – numpy, pandas, matplotlib, a sam model sieci neuronowej został napisany we frameworku Tensorflow w wersji 2.1.

# WSTĘP TEORETYCZNY

## Konserwacja predykcyjna

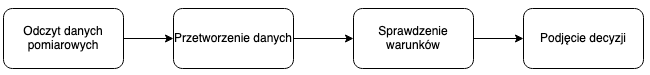
Wraz ze przyrostem obecności nowej technologii, konserwacja i utrzymanie ruchu odgrywają coraz ważniejszą rolę w wielu sektorach gospodarki. W elektroenergetyce niezawodność i poprawność działania są szczególnie istotne, ponieważ nawet drobna awaria może skutkować nieocenionymi stratami, a niekiedy i groźnymi w skutkach katastrofami. By temu zapobiec stan poszczególnych elementów systemu musi być regularnie badany. Wszystkie naprawy i konserwacje powinny być wykonywane na czas, uprzedzając awarie i w konsekwencji zniszczenie urządzeń i maszyn.

Głównym celem jest wypracowanie oszczędności związanych z wyeliminowaniem zbędnych interwencji serwisowych. Jednakże oprócz wymiernych korzyści materialnych, odpowiednia konserwacja pozwala na zwiększenie bezpieczeństwa, szczególnie w sytuacjach, gdzie awaria danego komponentu może skutkować zagrożeniem życia lub zdrowia.

Tradycyjna metoda konserwacji maszyn i urządzeń opiera się w dużej mierze na eliminowaniu zaistniałych usterek po fakcie. Zazwyczaj uzupełniana jest o wykonywanie przeglądów w ściśle określonych terminach, tworząc tym samym rodzaj konserwacji prewencyjnej (zapobiegawczej). Okresy te ustalane są wykorzystując wiedzę ekspercką i niekiedy modyfikowane są o stosowne adnotacje, jak skrócenie czasu pomiędzy kolejnymi przeglądami wraz z przebiegiem maszyny lub założenie, że po określonym czasie maszyna powinna zostać wymienioną na nową, czy przejść generalny remont. Określanie z góry ustalonych terminów konserwacji urządzeń, stosując nawet bardzo zaostrzone kryteria nie eliminuje występowania poważnych awarii. Ponadto konserwacja i utrzymanie ruchu stanowią znaczną cześć budżetu każdego przedsiębiorstwa i skrócenie okresów pomiędzy przeglądami lub bardziej restrykcyjne podejście w postaci częstszej wymiany lub regeneracji poszczególnych komponentów generuje dodatkowe koszty, które nie zawsze są do zaakceptowania z punktu widzenia biznesu.

Konserwacja predykcyjna (ang. predictive maintenance) to jeden z obszarów innowacji, który odgrywa coraz większą rolę w gałęziach sektora elektroenergetycznego, jak i całego przemysłu. Założeniem są systemy informatyczne, które określają stan techniczny danego urządzenia w czasie rzeczywistym. Wykorzystują w tym celu nieprzerwane pomiary jego parametrów. Na podstawie odpowiedniej kombinacji tych parametrów, sprawdzane są warunki określające konieczność wykonania konserwacji. Odpowiednio zaprojektowany i działający system konserwacji predykcyjnej sam zasygnalizuje wcześniej o zbliżającym się czasie wykonania niezbędnej naprawy urządzenia, zanim dojdzie do pogorszenia się jego parametrów lub całkowitego zniszczenia.

Klasyczna konserwacja prewencyjna skupiona jest na zaprojektowaniu najbardziej optymalnych odstępów czasowych pomiędzy wykonywanymi przeglądami i zakresami tych przeglądów. Konserwacja predykcyjna opiera się zaś na zaprojektowaniu optymalnych warunków, których spełnienie wygeneruje sygnał do konieczności wykonania konserwacji. Warunki oparte są na konkretnych kombinacjach wielkości parametrów urządzenia, które mierzone są w czasie rzeczywistym.



Rys 3.1. Schemat działania konserwacji predykcyjnej

Typowym przykładem zastosowania systemu konserwacji predykcyjnej w przemyśle jest monitorowanie dużych maszyn wirujących. Maszyna wyposażona jest w aparaturę pomiarową mierzącą jednocześnie wartości temperatur, drgań, poziomu hałasu, zawartości odpowiednich harmonicznych w prądzie, napięciu i innych parametrów. Analizowanie tych wartości w czasie rzeczywistym przez personel byłoby bardzo trudne do osiągnięcia. Co więcej, nawet zastosowanie systemu informatycznego porównującego te parametry często nie wystarcza. Zaimplementowanie logiki polegającej na reakcji na poszczególne poziomy graniczne mierzonych wielkości może nie dostarczyć zadowalających rezultatów lub co więcej powodować nadmierne fałszywe zadziałania (false positives). Nieliniowość niektórych parametrów i ich zmienna korelacja względem siebie w dynamicznych warunkach pracy maszyny jest bardzo trudna do zamodelowania stosując zwykłą logikę programistyczną lub klasyczne modelowanie statystyczne. Zaprojektowanie warunków, przy których system miałby wygenerować sygnał o zbliżającej się awarii, staje się w takim przypadku dużym wyzwaniem. Coraz częściej wykorzystuje się w tym celu sztuczne sieci neuronowe. Okazuje się, że potrafią one być bardzo dobrym analizatorem, szczególnie jeżeli należy poradzić sobie z problemem wielu zmiennych oraz zamodelować nietrywialne powiązania między nimi w czasie rzeczywistym.

## Sztuczne sieci neuronowe

Postęp jaki uczyniono w zakresie pozyskiwania danych oraz ich efektywnego przetwarzania spowodował silny rozwój dziedzin uczenia maszynowego, w tym sztucznych sieci neuronowych.

Sieci jako wyrafinowana technika modelowania, zdolna do odwzorowywania nawet nadzwyczaj złożonych i nieliniowych funkcji potrzebuje do swej pracy dużo większej ilości obserwacji niż popularne modele analityczne i statystyczne. Stąd teraz, w erze niemal nieograniczonego dostępu do danych i mocy obliczeniowych infrastruktur komputerowych metody oparte o sztuczne sieci neuronowe cieszą się coraz większą popularnością. Sieci neuronowe mogą być stosowane w praktycznie każdej sytuacji, gdzie pomiędzy zmiennymi zależnymi i niezależnymi istnieje rzeczywista zależność, nawet jeśli jest ona bardzo skomplikowana i niewyrażalna w klasyczny sposób.

Sieci umożliwiają także kontrolę nad złożonym problemem wielowymiarowości, który przy stosowaniu pospolitych metod statystycznych znacząco utrudnia lub nawet uniemożliwia próby modelowania funkcji z dużą ilością zmiennych niezależnych. Z tego też powodu nierzadko bywają fundamentem modeli opisujących bardzo złożone systemy jak elektrownie czy hale produkcyjne.

### Sieci neuronowe, a detekcja anomalii.

Obecnie znacząca większość rozwiązań oferowanych przez sztuczne sieci neuronowe wykorzystuje tak zwane uczenie nadzorowane (ang. supervised learning). Polega ono na dostarczeniu wraz z danymi uczącymi, wartości prawdziwych, które sieć ma przewidzieć. Jeśli na przykład sieć miałaby nauczyć się rozróżniać, czy na podanym na zdjęciu znajduje się kot czy pies, podczas fazy uczenia każdy obraz, oprócz jego numerycznej reprezentacji posiadałby także jedną z etykiet: pies lub kot. Sieć podczas uczenia znałaby prawdziwą etykietę, do której powinna się zbliżyć, aktualizując swoje wewnętrzne połączenia. Uczenie nadzorowane odgrywa w uczeniu maszynowym olbrzymią rolę. To tej metodzie zawdzięcza się postęp w przetwarzaniu mowy czy obrazów. Wymaga ona jednak dużej ilości wstępnie etykietowanych danych.

Wykorzystanie takiego podejścia w kontekście detekcji anomalii może być problematyczne. Jeśli sieć miałaby nauczyć się, które kombinacje danych wejściowych (natężenie prądu, temperatura, poziom hałasu itd.) odpowiadają awarii danego urządzenia, podczas uczenia należałoby przedstawić sieci takie reprezentatywne przypadki. To oznacza, że musielibyśmy posiadać zbiory danych, które reprezentują awarię konkretnych urządzeń. Uzyskanie takich danych, pomimo że możliwe, byłoby kosztowne i niepraktyczne. Nawet jeśli uwzględnimy przypadki, powstałe wskutek długo monitorowanego urządzenia, które po czasie uległo awarii okaże się, że takie dane, choć niewątpliwie przydatne, nie są tak łatwo dostępne. Awarie urządzeń w stosunku do czasu ich normalnej pracy są naturalnie dużo rzadszym zjawiskiem. Dane więc byłyby bardzo niezbalansowane – znacznie więcej próbek określających normalną prace, a tylko minimalna ilość próbek określających stan awarii. Problem niezbalansowanych danych znacznie utrudnia prawidłowe uczenie sieci.

Jak wspomniano, duże etykietowane zbiory danych monitorowanych maszyn są trudne do uzyskania, a proces tworzenia ich od podstaw jest bardzo kosztowny i wymaga dużej wiedzy eksperckiej. Z tego też powodu podczas wykrywania anomalii w pracy urządzeń w czasie rzeczywistym przy użyciu sztucznych sieci neuronowych nierzadko wykorzystuje się podejście nienadzorowane (ang. unsupervised learning).

Uczenie nienadzorowane zakłada brak obecności dokładnego, czy nawet przybliżonego wyjścia z danych uczących. Przytaczając przykład naszej maszyny wirującej, oprócz danych opisujących stan maszyny – poziom drgań, natężenie prądu itd. nie posiadamy żadnej informacji w jakim obecnie stanie znajduje się maszyna ani tym bardziej, jaki jest stopień jej zużycia czy degradacji poszczególnych komponentów. Dysponujemy tylko danymi zgromadzonymi podczas jej pracy.

Uczenie nienadzorowane jest bardzo specyficzną i trudną dziedziną uczenia maszynowego. Wymaga starannego doboru algorytmów, a wysnucie wniosków z efektów końcowych wymaga dużo więcej pracy i znajomości dziedziny rozwiązywanego problemu. Uczenie nienadzorowane skupia jednak coraz większą uwagę społeczności naukowej i stosowane jest coraz szerzej w świecie techniki i biznesu, ponieważ ta metoda potrafi doprowadzić do wniosków, których istnienia można by się w ogóle nie spodziewać analizując początkowo problem. W środowisku naukowym istnieje przekonanie, że algorytmy uczenia nienadzorowanego często potrafią odpowiedzieć na niezadane pytanie. Klasycznym przypadkiem jest analiza konsumencka wielkich marketów spożywczych. Koncerny spożywcze od lat stale monitorują preferencje grup konsumenckich podczas robienia zakupów wykorzystując do tego techniki uczenia nienadzorowanego. Wyjściem sieci analizującej całe grupy konsumentów i kupowanych przez nich produktów są ukryte, niewidzialne na pierwszy rzut oka zależności, które odpowiednio przeanalizowane doprowadzają do rozmieszczenia produktów na hali sklepowej w precyzyjnie przemyślany sposób.

Uczenie nienadzorowane w kontekście detekcji anomalii może być rozumiane jako sposób na znalezienie ukrytych informacji opisujących dane wejściowe – ich wzajemnej reprezentacji w przestrzeni względem siebie. Tak zamodelowane urządzenie, którego parametry monitorujemy w czasie rzeczywistym możemy potem rozpatrywać w kontekście ewentualnych zmian tych reprezentacji w przestrzeni – wystąpienia anomalii.

## Warystory tlenkowo-cynkowe

Ograniczniki przepięć zbudowane z warystorów tlenkowo-cynkowych stanowią istotny element systemu elektroenergetycznego. Ich rolą jest ograniczenie skutków przepięć zewnętrznych, jak i wewnętrzych (łączeniowych). Jednakże podczas ich awarii nierzadko dochodzi do zwarć czy wybuchów, które są niebezpieczne zarówno dla samego systemu, jak i personelu. Ponadto wskutek postępującej w czasie degradacji ich struktury wewnętrznej, mogą powodować nieplanowane przerwy w zasilaniu. Z tych powodów stan ograniczników przepięć musi być regularnie badany, aby nie dopuścić do ich niebezpiecznego w skutkach zniszczenia.

### Budowa

Warystory tlenkowo-cynkowe są ceramicznymi materiałami polikrystalicznymi charakteryzującymi się silnie nieliniową charakterystyką napięciowo-prądową. Procesy technologiczne wytwarzania ceramiki warystorowej powodują powstanie struktury składającej się z matrycy przewodzących ziaren ZnO otoczonych cienką warstwą międzyziarnową o właściwościach półprzewodzących. Tak ukształtowana mikrostruktura powoduje formowanie się podwójnych barier potencjału na złączach międzyziarnowych. Decydują one o powstaniu nieliniowego przewodnictwa elektrycznego struktury wewnętrznej warystora. [1 – kostyła]

Charakterystyka napięciowo-prądowa warystora ZnO opisana jest następującym równaniem:

Gdzie:

J – gęstość prądu,

U – napięcie,

c – stała

𝛂 – współczynnik nieliowości

### Diagnostyka

Degradacja struktury wewnętrznej warystora idzie w parze ze zmianą jego charakterystyki napięciowo-prądowej. Głównymi czynnikami powodującymi te zmiany to:

- wysoka temperatura,

- podwyższone napięcie pracy,

- wysoka wilgotność,

- piki przepięciowe

Często powyższe czynniki występują jednocześnie potęgując efekt degradacji.

Istnieje wiele technik diagnostyki ograniczników przepięć w czasie rzeczywistym. Poniżej wymieniono charakterystyki niektórych z nich.

1. Pomiar prądu.

Podstawowa metoda diagnostyki zakłada pomiar natężenia prądu płynącego przez warystor.

Pierwszym symptomem sugerującym początki degradacji warystora jest wzrost prądu płynącego przez jego strukturę przy określonym napięciu roboczym. Z punktu widzenia obwodu elektrycznego degradacja powoduje zmiany w wartościach rezystancji i pojemności. Sumaryczny prąd płynący przez warystor w znacznie większej mierze składa się jednak ze składowej pojemnościowej, której wielkość praktycznie nie ulega zmianie podczas degradacji. Metodę tę usprawnia się mierząc tylko składową stałą rezystancyjną prądu. [0 – kostyła] Ponadto pomiar prądu lub jego poszczególnych składowych powinien być także skorygowany o odpowiednie współczynniki poprawkowe, głównie temperaturowe i napięciowe [1 – olesz].

1. Pomiar temperatury

Wzrost prądu powoduje także wydzielanie się na warystorze większej ilości ciepła, co skutkuje podwyższeniem temperatury warystora. Tą zależność wykorzystuje się w metodzie opartej o pomiary temperatury. Do pomiarów wykorzystuje się czujniki bądź zaawansowane kamery termowizyjne.

1. Pomiar zawartości wyższych harmonicznych w prądzie.

Metoda ta uwzględnia zawartość trzeciej i dalszych harmonicznych w prądzie. Zawartość ich nie jest jednak zawsze jednakowo skorelowana ze wzrostem prądu, szczególnie w warunkach o podwyższonej wilgotności. Ponadto wartość może być zakłamana występowaniem wyższych harmonicznych w napięciu zasilającym [0 – kostyła].

1. Monitorowanie parametrów wewnętrznych warystora.

Metoda polega na wyznaczeniu parametrów wewnętrznych R i C warystora przed zużyciem, a następnie monitorowaniu tych parametrów w czasie pracy. Sygnał o postępującej degradacji wynika, ze zmiany tych parametrów względem wartości otrzymanych w warunkach nominalnych.

W celu wyznaczenia parametrów wewnętrznych warystora dokonuje się analizy budując ich schematy zastępcze. Dla napięć przemiennych przyjmuje się najczęściej połączone elementy R, C. Zaproponowany schemat elektryczny układu składa się z dwóch części adekwatnych do mikrostruktury warystora. Warstwa międzyziarnowa opisana jest jako schemat połączonych szeregowo-równolegle elementów R-C. Zjawiska występujące na granicy pomiędzy ziarnem ZnO i warstwy międzyziarnowej przedstawia schemat w postaci równoległego połączenia nieliniowej rezystancji i liniowej pojemności.

A close up of a logo

Description automatically generated

Rys 3.2. Model analityczny warystora ZnO [0 – kostyła]

Wyznaczenie parametrów może odbywać się drogą analityczną oraz przy użyciu sztucznych sieci neuronowych [0 – kostyła].

# ZAPROPONOWANE ROZWIĄZANIE

W pracy zaproponowano model oparty o sieć autoenkodera z warstwami rekurencyjnymi LSTM (Long Short Term Memory).

## Rekurencyjne sieci neuronowe

Klasyczne modele uczenia maszynowego i podstawowe struktury sieci neuronowych zakładają niezależność danych wejściowych. To znaczy, że konkretna próbka danych podawana na wejście modelu jest całkowicie niezależna od innych próbek występujących w zbiorze uczącym. Założenie to jest błędne w momencie, gdy dane reprezentowane są za pomocą szeregów czasowych. W przypadku gdy kolejność wystąpienia obserwacji w zbiorze ma znaczenie, potrzebne jest przetwarzanie danych w sekwencjach, czyli ciągach próbek o ściśle określonej kolejności.

Dobrze ilustrującym to przykładem są modele przetwarzania języka (ang. Natural Language Processing). Aby zrozumieć dobrze ludzki język, musimy znać kontekst w jakim wypowiadane są słowa. Czytając pojedyncze słowo lub wycinek tekstu z losowo umieszczonymi słowami trudno jest wysnuć odpowiednie wnioski. Model do przetwarzania ludzkiego języka musi być zatem uczony na całych sekwencjach słów, które występują w określonej kolejności.

W przypadku analizowania wielkości elektrycznych sytuacja jest podobna. Ciężko jest zbadać sygnał napięcia czy natężenia prądu analizując tylko pojedyncze próbki. Dopiero zamodelowanie całej sekwencji może pozwolić na trafniejszą analizę.

Rekurencyjne sieci neuronowe umożliwiają przetwarzanie danych sekwencjami, budując w swoich połączeniach coś na kształt wewnętrznej pamięci. Przyjmują jako wejście ciąg d-wymiarowych wektorów i przetwarzają je rekurencyjnie, wykonując te same operacje dla każdego elementu w czasie. Wyjście każdej próbki zależy nie tylko od zmiennych ją opisujących, ale także od jej poprzednich stanów. W ten sposób można zamodelować zależności, które uwidaczniają się dopiero po przeanalizowaniu całego ciągu danych w czasie, a nie pojedynczych próbek. Taka struktura umożliwia, aby informacja zawarta w pierwszej próbce miała wpływ nawet na ostatnią w sekwencji.

### Budowa

A picture containing clock

Description automatically generated

Rys. 4.1. Schemat rekurencyjnej sieci neuronowej.

Na powyższym rysunku przedstawiono uproszczony schemat działania rekurencyjnej sieci neuronowej. Sieć została symbolicznie rozwinięta do trzech sieci dla lepszego zilustrowania zachodzących operacji. W każdej chwili czasowej t na wejście sieci podawany jest wektor danych xt i stan ukryty H. Stan ukryty jest w istocie funkcją wartości z wszystkich dotychczas przetworzonych chwil w sekwencji. Macierz H aktualizowana jest po każdej chwili czasowej i jest współdzielona przez całą sieć. Każdy stan zależy więc od wszystkich poprzednich obliczeń.

Konsekwencją takiego przetwarzania danych jest wspomniana wcześniej struktura pamięci. Stany ukryte zawierają informacje bazujące na wcześniejszych krokach. Podczas treningu sieć dąży do takiej kombinacji parametrów ukrytych, aby wyjście było jak najbardziej zbliżone do przewidywanej wartości, uwzględniając przy tym sekwencyjny charakter danych. W praktyce sprowadza się to do takiego dostrojenia wag, aby były one aktualizowane nie tylko w oparciu o wektor wejściowy, ale także o poprzednie próbki w sekwencji. Może to przykładowo oznaczać, że próbka z chwili czasowej t-10 ma dużo większy wpływ na wynik niż pozostałe próbki, ponieważ zawiera kluczową dla predykcji informację. Sieć sama znajdzie w tym wypadku istotne elementy sekwencji i nada im odpowiednie wagi.

A close up of a sign

Description automatically generated

Rys. 4.2. Schemat rekurencyjnego neuronu – źródło [7]

<http://dprogrammer.org/rnn-lstm-gru>

W matematycznym ujęciu, wyjście w bieżącej chwili czasowej składa się z dodania wektorów - wyjściowego z poprzedniej chwili t-1 i wejściowego z bieżącej chwili t. Całość przepuszczana jest przez nieliniową funkcję aktywacji tanh (tangens hiperboliczny), która sprowadza wartości do przedziału -1; 1.

### Wyjście sieci rekurencyjnej

W podstawowej konfiguracji sieć przetwarza całą sekwencję danych i generuje na swoim wyjściu sekwencje o tej samej długości. Wyjście jest generowane dla każdej próbki w czasie, więc gdy przykładowo na wejście zostanie podana sekwencja składająca się z 3 próbek, wyjście także będzie ciągiem 3 próbek. Nie jest to jedyna możliwa konfiguracja. Sieci rekurencyjne mogą przetwarzać dane na wiele sposobów. Gdy modelują na przykład klasyfikator wyjściem może być jedna próbka. Ponadto podawane na wejście sekwencje mogą być różnej długości. Schematy niektórych konfiguracji pokazano na rysunku 4.3.

A picture containing drawing, clock

Description automatically generated

Rys. 4.3. Konfiguracje pracy sieci rekurencyjnych – źródło [7]

<http://dprogrammer.org/rnn-lstm-gru>

### Problem zanikających gradientów.

Trening sieci neuronowych składa się z następujących etapów:

- przejście wektora danych wejściowych przez sieć (forward pass),

- porównanie wyjścia z sieci (predykcji) z wartością prawdziwą – obliczenie błędu sieci,

- obliczenie gradientów parametrów sieci względem otrzymanego błędu przy użyciu algorytmu wstecznej propagacji,

- korekcja parametrów sieci w oparciu o obliczone gradienty – zmniejszenie błędu sieci.

Sieci rekurencyjne mogą być znacznie trudniejsze do trenowania niż klasyczne sieci neuronowe. Sekwencyjny charakter przetwarzania danych niesie ze sobą dodatkowy koszt – konieczność uwzględnienia każdej chwili czasowej również podczas treningu. To powoduje, że ilość operacji obliczeniowych niezbędnych do wykonania w celu propagacji wstecznej błędu rośnie wraz z długością sekwencji. Z tego powodu sieci rekurencyjne, podobnie jak bardzo głębokie i złożone sieci neuronowe cierpią na przypadłość zanikających gradientów.

Sieć neuronowa jest w istocie bardzo złożoną funkcją matematyczną składającą się z wielu podfunkcji. Obliczenie gradientów jej poszczególnych parametrów w celu późniejszej korekcji realizowane jest przy użyciu znanej w matematyce reguły łańcuchowej. Łatwo wyobrazić sobie, że parametr występujący w sieci wcześniej – bliżej warstwy wejściowej, a dalej wyjściowej, będzie podczas kalkulacji gradientu obarczony większą ilością operacji matematycznych do wykonania. Jeśli dodatkowo poszczególne składowe tych operacji będą miały wartość mniejszą od 1, finalnie wynik dążył będzie do 0.

Z tego też powodu przy głębokich sieciach, czyli takich, w których dane wejściowe są przetwarzane przez wiele warstw zanim dojdą do neuronów wyjściowych, istnieje ryzyko, że parametry pierwszych warstw zostaną zignorowane, ponieważ ich gradient, obliczony względem błędu sieci będzie miał wartość bliską 0.

Poniższy rysunek poglądowo przedstawia problem zanikających gradientów w głębokich sieciach neuronowych.

A picture containing knife, table

Description automatically generated

Rys. 4.4. Zanikające gradienty w głębokiej sieci neuronowej – źródło [7]

https://trendy00develope.tistory.com/37

Jak widać, pierwsze warstwy sieci podczas operacji wstecznej propagacji błędu zostają wygaszone, czyli ich gradienty są bliskie zeru. Podczas treningu nie zostają one więc korygowane, wskutek czego pozostają dla sieci niewykorzystaną przestrzenią do modelowania informacji. W ten sposób sieć, podczas swojej pracy weźmie pod uwagę tylko ostanie, najbliższe wyjściu warstwy.

W sieciach rekurencyjnych iteracyjne przejście przez całą sekwencję powoduje ten sam efekt – zanikanie gradientów w początkowych próbkach. Skutkuje to znacznym ograniczeniem wykorzystania takich sieci przy pracy z długimi sekwencjami. Gradienty we wcześniejszych stanach stają się bardzo małe i znika możliwość zachowania historii tych stanów. Sieć traci w tym wypadku swoje zdolności do zapamiętania i modelowania sekwencji z uwzględnieniem wszystkich, także początkowych próbek. Propagacja wsteczna w czasie jest zbyt wrażliwa na niedawne zakłócenia. Choć w teorii proste sieci rekurencyjne są w stanie uczyć się zależności długoterminowych, w praktyce ograniczają się tylko do zależności krótkoterminowych. Z tego powodu powstały ich bardziej rozbudowane odpowiedniki.

### Sieci LSTM

Sieci LSTM (ang. Long Short Term Memory) to sieci rekurencyjne, które powstały w celu wyeliminowania problemu zanikającego gradientu oraz poprawy pracy z długimi sekwencjami danych. Względem klasycznej sieci neuronowej, ich struktura została uzupełniona o tzw. bramki oraz dodatkowy stan ukryty przechowujący pamięć długoterminową.

A close up of a clock

Description automatically generated

Rys. 4.5. Komorka sieci LSTM – źródło [7] – zamienić na obrazek z książki

<http://dprogrammer.org/rnn-lstm-gru>

Komórka sieci LSTM jest bardziej skomplikowaną strukturą. Zawiera ona w sobie dodatkowy stan ukryty, oznaczony literą C oraz bramki. Stan ukryty jest strukturą przechowującą długoterminowe zależności. Informacje w nim zawarte mogą być dynamicznie zapisywane i usuwane. Jeśli nie ma ingerencji z zewnątrz, stan komórki pozostaje bez zmian. Jest to element zapewniający sieci pamięć. Bramki są zaś swojego rodzaju zaworami. Regulują przepływ informacji w sieci, umożliwiając stanowi ukrytemu zachowanie najcenniejszych informacji w sekwencji.

Typowa sieć LSTM zawiera 3 bramki:

- zapominania (ang. forget gate), oznaczona na rysunku literą f,

- wejścia (ang. input gate), oznaczona literą i,

- wyjścia (ang. output gate) oznaczona literą o

Bramka zapominania swoją nazwę zawdzięcza temu, że decyduje o ilości informacji, które mają zostać wykasowane ze stanu ukrytego. Opiera swoją decyzję na poprzednim wyjściu ht-1 i bieżącym wejściu xt. Łączy te informację i kompresuje za pomocą funkcji logistycznej, a następnie przemnaża otrzymane wartości z wektorem stanu ukrytego. Dzięki temu, że funkcja logistyczna zwraca wartości z przedziału 0; 1 informację mogą być całkowicie zachowane – mnożenie przez 1 lub zapomniane - mnożenie przez 0. Oznacza to, że pamięć LSTM może się samoistnie pozbyć niepotrzebnych informacji z wektoru stanu ukrytego.

Bramka wejściowa decyduje o tym, jakie nowe informacje zostaną dodane do komórki pamięci. Odbywa się to w dwóch etapach. W pierwszym podejmowana jest decyzja o tym, czy informacje w ogóle zostaną dodane. Tak, jak w przypadku bramki zapominania, decyzja bazuje na wartościach ht-1 i xt.. Funkcja zwraca wynik 0 lub 1 za pośrednictwem funkcji logistycznej dostępnej dla każdego bloku wektora. W rezultacie sieć LSTM, może dodawać do swojej pamięci, tylko specyficzne informację.

Wektor informacji, który został zakwalifikowany do dodania jest obliczany na podstawie poprzedniego wyjścia ht-1 i skompresowany za pomocą funkcji tanh

Ostatecznie bramka zapominania i bramka wejściowa decydują o zawartości stanu ukrytego – pamięci sieci.

Ostatnia bramka decyduje z czego składać ma się wyjście w bieżącej chwili czasowej. Pobiera ona wartości ht-1 i xt. jako swoje wejście i zwraca wektor liczb z przedziału 0; 1 za pomocą funkcji logistycznej. Wyjście 0 oznacza, że blok komórki nie zwraca w bieżącej chwili żadnej informacji, a 1 że na wyjście przekazywana jest pełna informacja z bieżącej próbki.

Ostateczna wartość wyjścia to kombinacja otrzymanego powyżej wyjścia z bieżącej komórki oraz dotychczas zarejestrowanego stanu ukrytego – pamięci długotrwałej sieci.

Obrazek połączonych sekwencyjnie neuronów

Ponieważ wszystkie te wielkości są dostępne na zewnątrz komórki, tak jak w przypadku klasycznych neuronów rekurencyjnych, istnieje możliwość ich połączenia, aby przetwarzały całe sekwencje danych i trenowania całej sieci za pomocą wstecznej propagacji w czasie.

Dodatkowy stan ukryty jest niezależnym wektorem, magazynującym dotychczasowe informacje. Tylko bramka zapominania jest w stanie wymazać z niech pamięć, co czyni sieć LSTM zdolną do przechowywania długoterminowych zależności, a informację w niej zawarte mogą pozostać niezmienione przez długi czas.

### Sieci dwukierunkowe LSTM

Sieci rekurencyjne umożliwiają także pracę w trybie dwukierunkowym. Polega to na zdublowaniu warstwy komórek rekurencyjnych, tak aby jedna przetwarzała sekwencje danych licząc od pierwszej próbki do ostatniej, a druga od ostatniej do pierwszej. Istnieją dwa stany ukryte w danej chwili czasowej otrzymane w wyniku tych dwóch iteracji. Ostateczne wyjście składa się z odpowiedniej kombinacji tych dwóch stanów ukrytych – najczęściej połączenia (konkatenacji), lecz niekiedy ich sumy lub średniej.

A close up of a clock

Description automatically generated

Rys. 4.6. Sieć rekurencyjna dwukierunkowa – źródło [7]

<https://colah.github.io/posts/2015-09-NN-Types-FP/>

Sieci dwukierunkowe zostały stworzone z myślą o zwiększeniu ilości informacji dostępnej dla sieci podczas przetwarzania sekwencji. Sieć w danej chwili czasowej ma do dyspozycji informację przetworzone od początku sekwencji, a także od jej końca. Jest to szczególnie istotne, w przypadkach, kiedy modelowane zjawisko nie jest zależne tylko od przeszłych próbek, ale także od przyszłych. Uzyskana w ten sposób informacja stanowi swojego rodzaju kontekst, dostępny w każdej chwili czasowej. To rozwiązanie dobrze sprawdza się przy tłumaczeniu języka – często ostatnie słowa mają kluczowy wpływ na znaczenie całego zdania. Przetwarzanie sekwencji słów jednocześnie od początku i od końca może pozwolić na trafniejszą predykcję w danej chwili czasowej.

Sieci dwukierunkowe są kosztowniejsze w użyciu – wymagają większych zasobów obliczeniowych. Ponadto do ich trenowania potrzebne są kompletne sekwencje danych. Ponieważ w danej chwili czasowej sieć potrzebuje informacji z dwóch stron sekwencji, niezbędne jest, aby sekwencja była kompletna i zawierała informację na obu końcach (?)

## Autoenkodery

Autoenkodery są symetrycznymi sieciami neuronowymi trenowanymi w sposób nienadzorowany. Ich zadaniem jest odtworzenie danych wejściowych. Składają się z dwóch części: enkodera i dekodera.

Enkoder jest funkcją, która rzutuje dane wejściowe na ich utajoną reprezentację. Przyjmuje postać:

Dekoder zaś jest funkcją, która odtwarza z tej utajonej reprezentacji dane wejściowe.

, gdzie:

s – nieliniowa funkcja aktywacji

A close up of a map

Description automatically generated

Rys. 4.7. Schemat ideowy autoenkodera – źródło [7]

https://blog.paperspace.com/autoencoder-image-compression-keras/

Warstwa wyjściowa ma ten sam rozmiar co wejściowa, ponieważ jej celem jest zrekonstruowane własnych wejść, a nie przewidywanie wartości docelowych. Wyjście autoenkodera jest zatem przybliżeniem jego wejścia. Symbolicznie można opisać w następujący sposób:

Trenowanie autoenkodera sprowadza się do zminimalizowania błędu odwzorowania wektora wejściowego. Funkcję strat opisuje błąd rekonstrukcji – najczęściej błąd średniokwadratowy (ang. Mean Squared Error – MSE), który definiuje jak dobrze sieć zrekonstruowała podany na wejście wektor danych. Ta reprezentacja zostaje zapisana jako parametry warstwy utajonej „Code”, która ma najczęściej mniejszy wymiar od warstwy wejściowej. Taka konfiguracja zmusza sieć do znalezienia najbardziej skutecznej i kompaktowej reprezentacji danych wejściowych z minimalną utratą informacji. Głębsze autoenkodery, w których wykorzystane są nieliniowe funkcje aktywacji bardzo dobrze radzą sobie ze znalezieniem ukrytych lub skompresowanych cech danych wejściowych.

Struktury autoenkoderów wykorzystuje się w wielu dziedzinach, m.in. podczas detekcji anomalii, kompresji wymiarów danych wejściowych czy odszumianiu obrazów. Z powodu zdolności do wyuczenia ukrytych rozkładów danych wejściowych znalazły one także zastosowanie w modelach generatywnych transferujących style malarskie lub generujące utwory muzyczne. Wykorzystuje się je także w kontrowersyjnym projekcie o nazwie DeepFake, który polega na zamianie twarzy osób występujących w filmach bądź na zdjęciach.

## Opis zaprojektowanego algorytmu

Zaproponowane rozwiązanie do wykrycia anomalii w szeregach czasowych reprezentujących próbki napięcia i prądu warystora wykorzystuje architekturę autoenkodera z dwukierunkowymi warstwami rekurencyjnymi LSTM. Schemat działania takiego rozwiązania jest następujący:

Krok 1.

Identyfikacja „prawidłowego” zbioru danych. Chodzi o próbki, co do których mamy pewność, że reprezentują encje niebędące anomaliami. Identyfikacja, ogólnie rzecz biorąc bazuje na danych historycznych oraz na założeniu, że oficjalnie nie rozpoznano żadnych anomalii.

Takie podejście jest często określane mianem uczenia seminadzorowanego. Nie jest ono czysto nienadzorowane. Bazuje ono bowiem na założeniu, że większość obserwacji jest pozbawiona anomalii.

Krok 2.

Trening modelu na sekwencjach „prawidłowych” danych. Zadaniem modelu jest wyuczenie się rozkładu danych odpowiadającemu prawidłowej pracy warystora – przed pojawieniem się oznak degradacji struktury wewnętrznej. Autoenkoder powinien z możliwie jak najmniejszym błędem odtwarzać podane na wejście treningowe sekwencje danych.

Krok 3.

Zaprzestanie trenowania modelu i przełączenie w tryb czuwania. Tryb czuwania polega na podawaniu sieci rzeczywistych danych i sprawdzaniu otrzymanego błędu rekonstrukcji autoenkodera. Jeśli błąd mieści się w przedziale, który wcześniej odnotowano na danych uczących – sekwencja danych uznawana jest za prawidłową. Jeśli zaś błąd jest większy od przyjętego, czyli sekwencja zawiera inny rozkład zmiennych niż sieć widziała dotychczas – traktujemy to jako anomalię. Dzięki wytrenowaniu modelu na danych nie zawierających anomalii istnieje duże prawdopodobieństwo poprawnej rekonstrukcji danych reprezentujących dobry stan warystora, ale w przypadku sekwencji odstających błąd rekonstrukcji wzrasta ponad nieodnotowaną wcześniej wartość i istnieje możliwość zasygnalizowania pojawienia się degradacji urządzenia.

# PRAKTYCZNA IMPLEMENTACJA

Niniejszy rozdział poświęcony jest praktycznemu aspektowi stworzenia modelu do detekcji anomalii w warystorze tlenkowo-cynkowym. Całość została napisana w wysokopoziomowym języku programistycznym Python w wersji 3.8 i jest dostępna pod następującym linkiem:

https://github.com/r-zareba/praca\_inzynierska

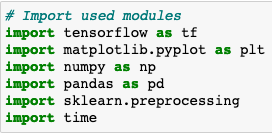
## Wczytanie i przygotowanie danych uczących

Jak każdy model uczenia maszynowego, sieć neuronowa do treningu potrzebuje danych. Jako zbiór uczący wykorzystano zarejestrowane próbki napięcia i natężenia prądu warystora podczas jego pracy w systemie elektroenergetycznym.

Ze względu na złożoność obliczeniową zaproponowanej architektury sieci oraz ograniczone zasoby komputerowe dostępne podczas tworzenia niniejszej pracy, zdecydowano, że dane uczące składać się będą jedynie z 10 tysięcy próbek. Zatem wektor danych uczących będzie miał wymiar (10000, 2). Składają się na niego dwa pełne okresy przebiegów napięcia i prądu.

Praca ma charakter eksperymentalny, więc dla uproszczenia założono, że podane przebiegi występowały przez dłuższy cały czas pracy warystora.

Przed rozpoczęciem pracy ze środowiskiem programistycznym, wczytano niezbędne moduły:



Rys. 5.1. Wczytanie użytych w projekcie modułów (bibliotek).

Dane wczytano z pliku o rozszerzeniu csv (ang. comma seperated values), a następnie oczyszczono i dokonano skalowania do przedziału (0, 1) wykorzystując do tego odpowiednie funkcje.

A screenshot of a social media post

Description automatically generated

Rys. 5.2. Wczytanie i przygotowanie danych treningowych

Dane narysowano na wykresie, korzystając z biblioteki matplotlib.

A close up of a map

Description automatically generated

Rys. 5.3. Wykresy danych uczących – napięcia i natężenia prądu.

Następnym krokiem było stworzenie odpowiednich tensorów treningowych. Sieć rekurencyjna działa na sekwencjach danych. Dotychczasowe dane zawierały dwa wymiary (10000, 2). Aby sieć mogła dostawać na wejście całe sekwencje, potrzebne było stworzenie dodatkowego wymiaru.

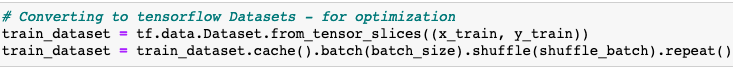
A screenshot of a cell phone

Description automatically generated

Rys. 5.4. Stworzenie tensorów treningowego zawierającego sekwencje próbek.

Jak widać, stworzone tensory mają wymiar (9968, 32, 2), co można interpretować jako 9968 sekwencji, każda będąca ciągiem 32 występujących po sobie próbek, a dodatkowo każda próbka opisana jest przez 2 zmienne (natężenie prądu i napięcie). Warto zwrócić uwagę na miejsce definicji tensora „y\_train”. Zawiera on elementy, które sieć powinna przewidzieć. Jest on wykorzystywany podczas treningu, do obliczenia funkcji błędu. Z racji tego, że celem autoenkodera jest rekonstrukcja jego własnego wejścia, tensor wyjściowy jest więc tym samym co wejściowy. Programistycznie rozwiązano to za pomocą operacji prostego przypisania.

Końcowym etapem jest stworzenie finalnych zbiorów uczących za pomocą obiektów biblioteki tensorflow. Ten krok ma na celu jedynie zoptymalizowanie operacji na danych podczas uczenia sieci i jest opcjonalny.



## Uczenie

## Walidacja

# WNIOSKI