**Damian Kobyliński**

Analiza zużycia wybranego elementu w oparciu o metodę zubożania cech oraz rekordów

Zadanie Projektowe

Opiekun pracy:  
dr inż. Marek Bolanowski

Rzeszów, 2023

Spis treści

[1. Wprowadzenie 5](#__RefHeading___Toc682_3979592501)

[2. Przedstawienie projektu 6](#__RefHeading___Toc684_3979592501)

[2.1. Biblioteki 6](#__RefHeading___Toc695_3979592501)

[2.2. Baza danych 6](#__RefHeading___Toc697_3979592501)

[2.3. Mapa koleracji pomiędzy cechami 9](#__RefHeading___Toc699_3979592501)

[2.4. Podział na grupę trenigową oraz testową 10](#__RefHeading___Toc701_3979592501)

[2.5. Funkcje obliczania Fscore 10](#__RefHeading___Toc703_3979592501)

[2.6. Model sieci uczącej 11](#__RefHeading___Toc705_3979592501)

[2.7. Trenowanie sieci 12](#__RefHeading___Toc707_3979592501)

[2.8. Prezentacja danych szczegółowych 12](#__RefHeading___Toc709_3979592501)

[2.9. Dokładność trenowanego modelu 16](#__RefHeading___Toc711_3979592501)

[2.10. Prezentacja wyników i wnioski 16](#__RefHeading___Toc713_3979592501)

[3. Podsumowanie 18](#__RefHeading___Toc686_3979592501)

[Załączniki 19](#__RefHeading___Toc688_3979592501)

[Literatura 20](#__RefHeading___Toc690_3979592501)

# Wprowadzenie

W pozniższym sprawozdaniu projektowym chciałbym opisać zadanie projektowe oparte na analize zurzycia wybranego elementu w oparciu o metodę zuborzania cech jak i ilosci rekordów badanego elementu. Jako baza danych do analizy wykorzystałem dane z Instytucji Energii Naturalnej na 14-stu bateriach NMC-LCO 18650                  z pojemnością 2.8 Ah. Baterie zostaly przetestowane na 1000 cyklach w temperaturze 25oC.

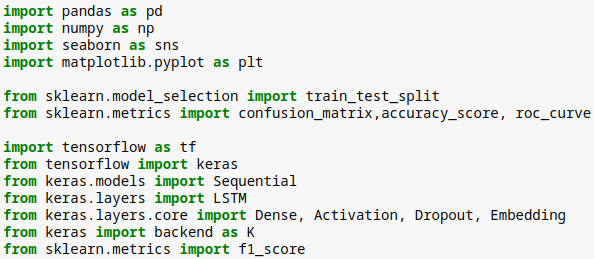
Projekt wykonany głównie w celu zaznajomienia z podstawowymi funkcjami języka programowania Pythona. Dodatkowo zostały wykorzystane dodatkowe biblioteki:

* Pandas – biblioteka do manipulacji i analizy danych,
* Numpy - biblioteka do manipulacji i analizy danych,
* Seaborn – biblioteka pozwalajaca zwizualizować informacje o danej bazie danych,
* Matplotlib – biblioteka służąca do rysowania wykresów i wizualizacji danych,
* Sklearn – biblioteka zawierająca pakiety funkcji do uczenia maszynowego,
* Tensorflow – biblioteka służąca do obliczeń numerycznych i tworzenia modeli uczenia maszynowego,
* Keras – wysokopoziomowa biblioteka uczenia maszynowego,

# Przedstawienie projektu

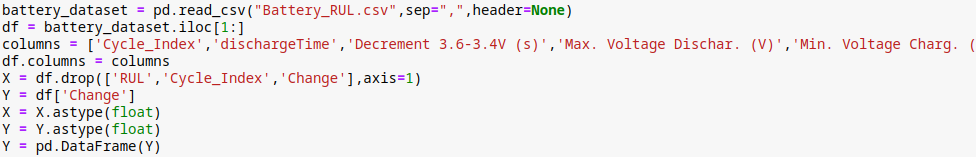
W tym rozdziale opisany został cały szkielet projektu.

## Biblioteki

Zdjęcie 1: Przedstawienie wykorzystanych bibliotek

## Baza danych

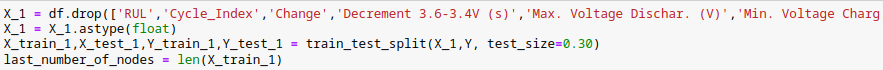
Wykorzystana baza danych:

Zdjęcie 2: Przedstawienie bazy danych

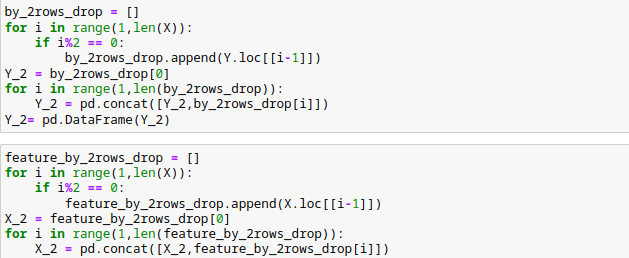
W przedstawionym powyżej kodzie, wyczytywany jest plik typu CSV. Tworzona jest kopia tego pliku z usunięciem pierwszego wiersza w którym zawarte są nazwy kolumn – je umieszczamy ręcznię. Następnie dzielona jest baza danych na kolumny z zawartymi cechami jak i z kolumnami labelów.

W trakcie zubażania zarówno kolumn z cechami, jak i ilości rekordów, zmienia się ilość danych.

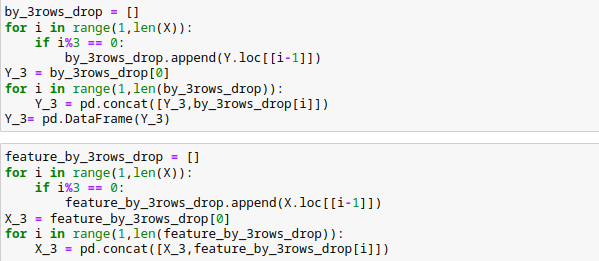
* Zubażanie co 2 kolumne z cechami
* Zubażanie do jednej kolumny z cechami

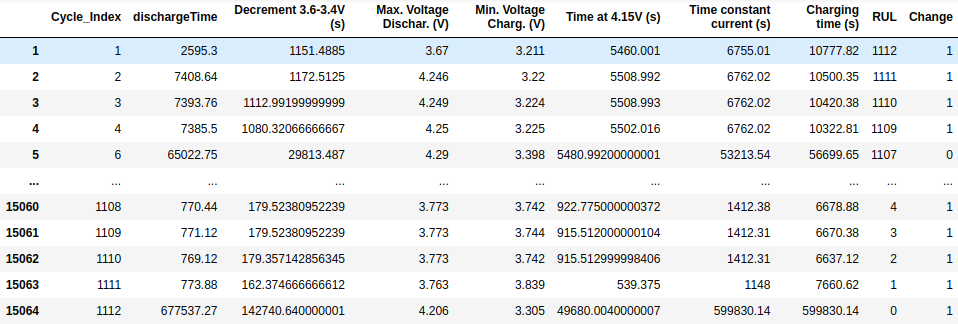


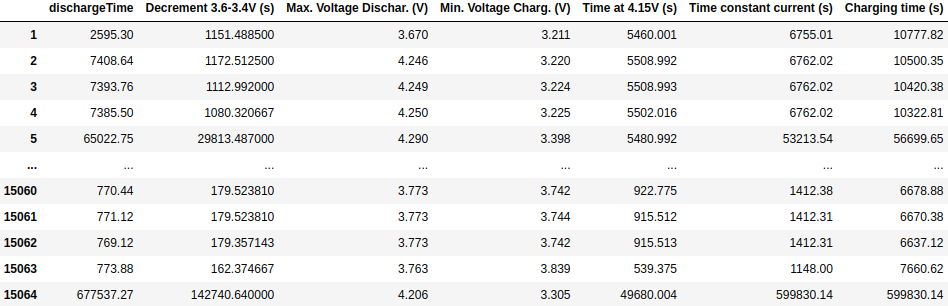
* Zubożanie co o połowe liczby rekordów w bazie przez usunięcię co 2 rekordu



* Zubożanie do 1/3 liczby rekordów w bazie przez pozostawienie co 3 rekordu



Zdjęcie 3: Cała baza danych

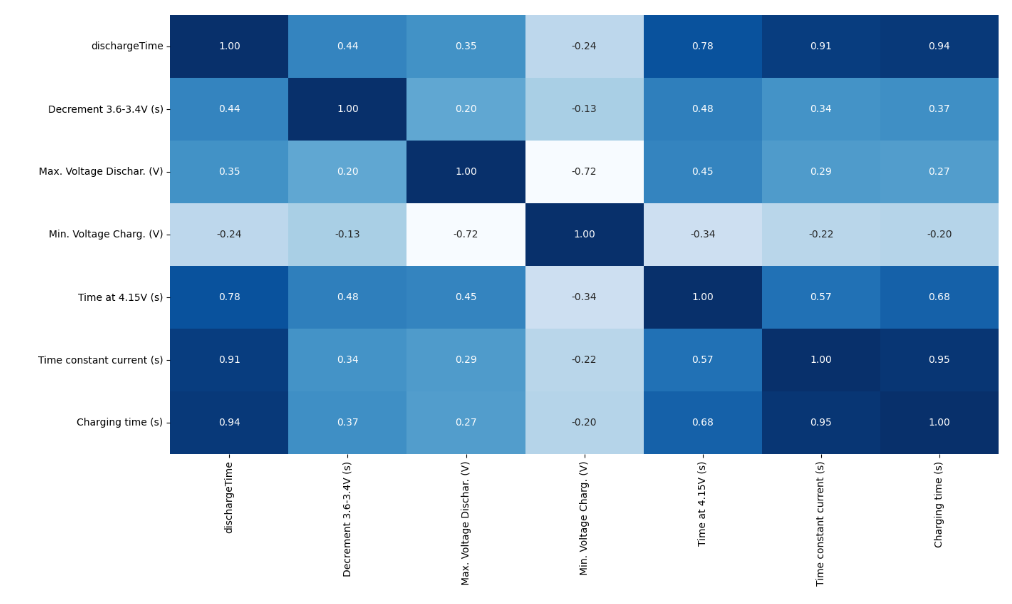
Zdjęcie 4: Kolumny cech

Zdjęcie 5: Kolumna labeli

## Mapa koleracji pomiędzy cechami

Przedstawienie korelacji pomiędzy poszczególnymi kolumnami cech, wykorzystanie do tego funkcji zawartych w bibliotece seaborn.

Zdjęcie 6: Kod przedstawiający mapę korelacji tabeli z kolumnami cech

Zdjęcie 7: Mapa korelacji cech

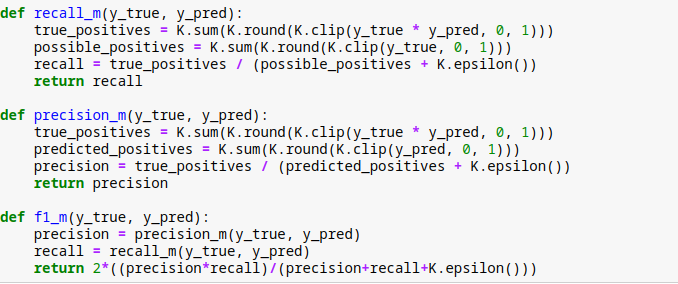
## Podział na grupę trenigową oraz testową

Grupa testowa to 30% rekordów całej bazy danych.

Zdjęcie 8: Podział na grupę treningową oraz testową

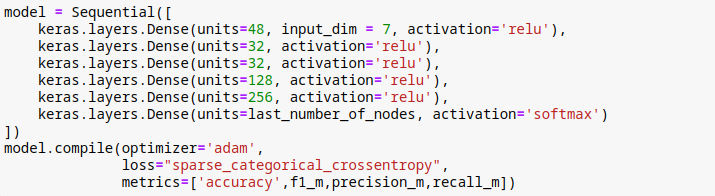
## Funkcje obliczania Fscore

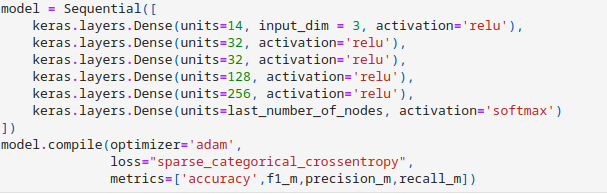
Funkcje pozwalające na obliczenie i uwzględnienie metryki fscore w trenowanym modelu uczenia maszynowego.

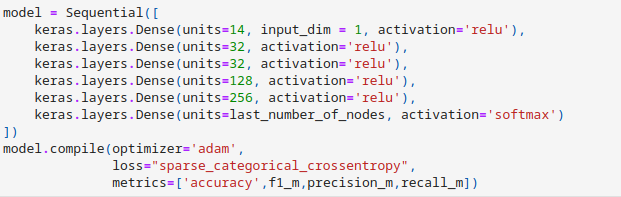
Zdjęcie 9: Funkcje pozwalające na obliczenie metryki FScore

## Model sieci uczącej

Prezentowany model posiada 6 warstw aktywacyjnych – 5 typu „relu”, 1 „softmax”. Pierwsza warstwa posiada w zależności od ilości kolumn z cechami, taką ilość wejść, natomiast posiada taką samą ilość wyjść wynikowych. Model optymalizacji został wybrany „Adam” oraz oprócz metryki Fscore, obliczana jest dokładność modelu uczącego.

Zdjęcie 10: Model sieci uczącej z informacji o jej kompilacji

Zdjęcie 11: Model sieci uczącej, zbioru danych zubażałych o co 2 kolumne

Zdjęcie 12: Model sieci uczącej, zbioru danych zubożałych do 1 wartościowej kolumny

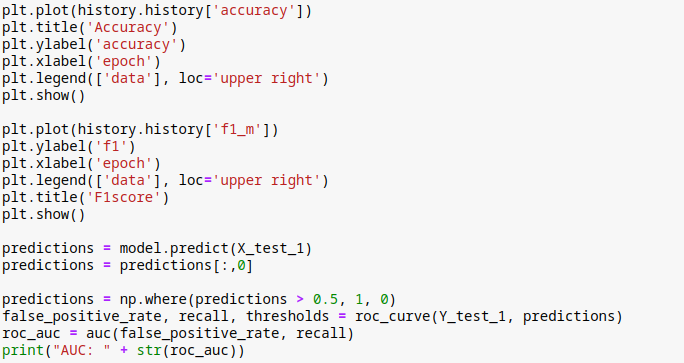
## Trenowanie sieci

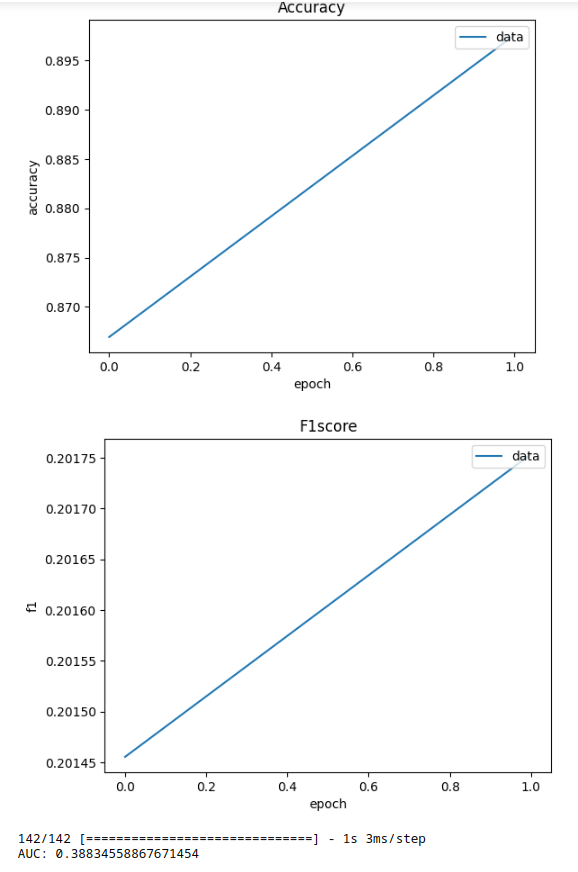
Trenowanie sieci odbywa się w 2 epokach.

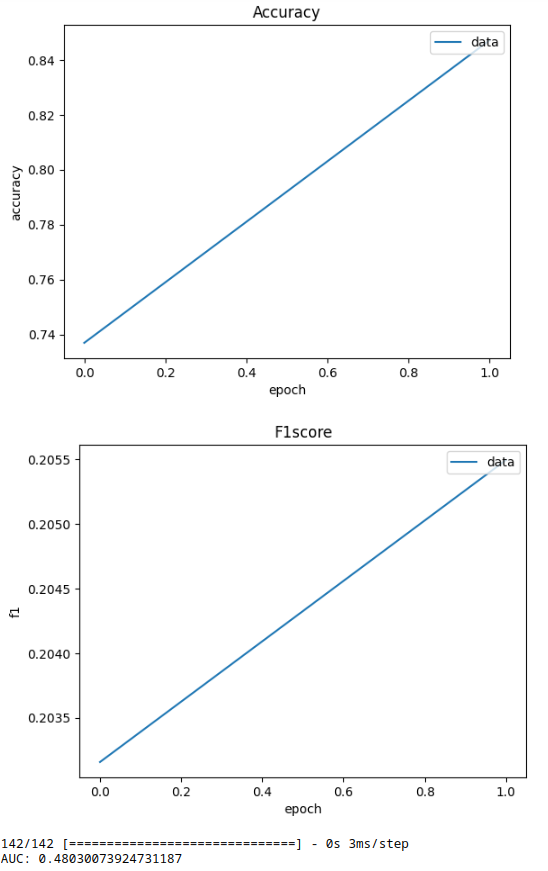
Zdjęcie 13: Trenowanie modelu sieci

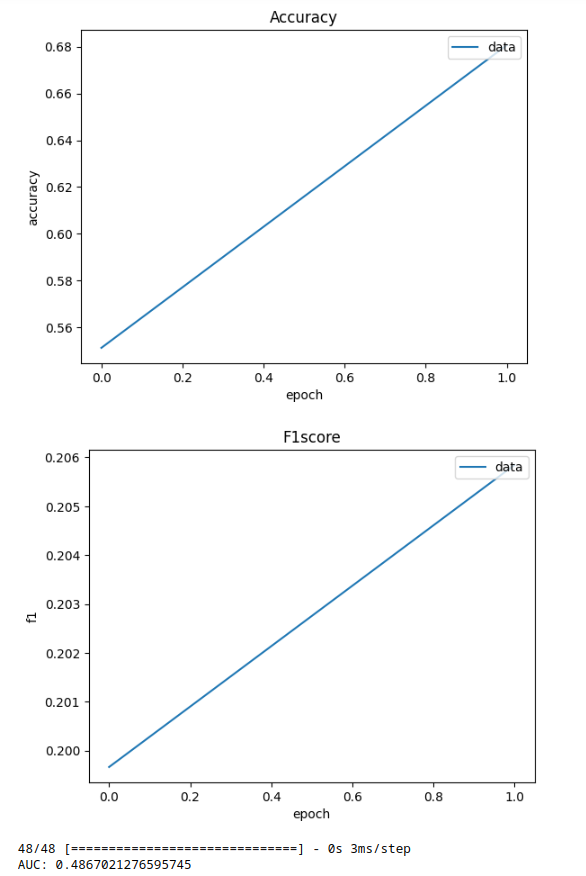
## Prezentacja danych szczegółowych

Prezentacja danych do każdej iteracji zubożania danych ( do każdej zmiany ilości kolumn z cechami).

Zdjęcie 14: Prezentacja danych szczegółowych

Zdjęcie 15: Wykresy szczegółowe wyniku modelu sieci uczącej dla kolumn z zawartymi wszystkimi cechami

Zdjęcie 16: Wykresy szczegółowe wyniku modelu sieci uczącej dla co drugiej kolumny cech

Zdjęcie 17: Wykresy szczegółowe wyniku modelu sieci uczącej dla jednej kolumny z cechami

## Dokładność trenowanego modelu







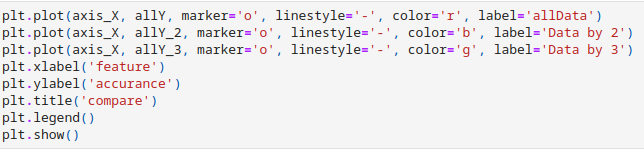
Do powyższych wartości średnich metryki dokładności modelu, utworzony jest wykres przedstawiający ogólny wynik wszystkich iteracji zubażania danych z których możemy wysnuć wnioski odnośnie ile kolumn cech oraz danych potrzebujemy minimalnie, aby stwierdzić, że dana bateria może być jeszcze do użytku.

## Prezentacja wyników i wnioski

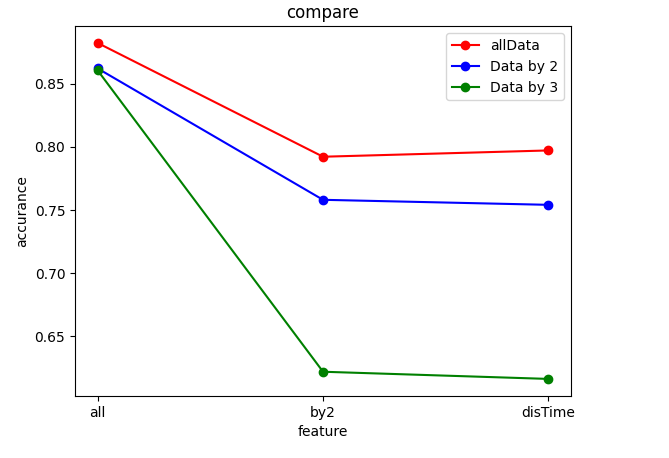








Powyższy kod pozwala wykonać dany wykres:



W przedstawionym powyżej wykresie widać, że najlepszą dokładność w proporcji do ilości danych widać gdy kolumny z cechami zostały zredukowane co 2 jak i ilość rekordów została zredukowana co druga.

# Podsumowanie

Przedstawiony powyżej projekt zawierał podstawowe zapoznanie z ekstracją oraz działaniem nad danymi w języku Python, jak i zawierał w sobię część wiedzy teorytycznej jak i praktycznej uczenia maszynowego. Dalszy rozwój i praca nad projektem może skutkować uwzględnienie jej jako pracę inżynierską, a jej wyniki byłyby możliwe do wykorzystania w życiu codziennym np. w magazynach.

# Załączniki

Praca umieszczona została na platformie GitHub wraz z plikiem HTML prezentujący cały kod projektu wykonany na jupyter-notebook.

# Literatura

[1] https://bolanowski.v.prz.edu.pl/download/word-template-2.html

[2] <https://stackoverflow.com/>-rozwiązywanie problemów technicznych

[3] https://towardsdatascience.com/building-our-first-neural-network-in-keras-bdc8abbc17f5