

Derin öğrenme tabanlı Elektrikli Ev Aletleri Veri Setinin Sınıflandırılması

Gürkan KÜÇÜKYILDIZ*

¹Uşak Üniversitesi, Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği Bölümü, Turkey
*e-mail: gurkan.kucukyildiz@usak.edu.tr, ORCID: 0000-0003-2744-0666

Geliş Tarihi:27.04.2023 ; Kabul Tarihi:02.06.2023

Öz

Anahtar kelimeler

LSTM, Bi-LSTM, ev aletleri, akım/gerilim, genetik algoritma

Elektriğe bağlı olan her ev aletinin akım/gerilim karakteristiği farklı olduğundan, bu cihazların her birinin şebekeden çektiği gücün özelliği farklı olmaktadır. Bu nedenle şebekeye bağlı olan cihazın tipinin tespiti cihazın şebekeden çektiği harmoniğin tespit edilmesinde ve de düzeltilmesinde önemli rol oynamaktadır. Bu çalışma kapsamında farklı derin öğrenme teknikleri kullanılarak "ACS-F2 Elektrikli Ev Aletleri Veri Seti" üzerinde sınıflandırma gerçekleştirilmiştir. ACS-F2 veri setinde toplamda 15 farklı sınıf için 225 cihaz bulunmasına karşın, çalışma kapsamında yapılan ön işlemler ile veri setindeki sınıf sayısı 14'e indirilmiştir. Sonrasında LSTM, FeedForwardNet, çift yönlü LSTM(Bi-LSTM) ve parametreleri genetik algoritma tarafından optimize edilmiş Bi-LSTM kullanılarak sınıflandırma yapılarak sınıflandırıcının performansları karşılaştırılmıştır. Yapılan çalışma kapsamında parametreleri optimize edilmiş sınıflandırıcının diğer yöntemlerden daha başarılı sonuçlar elde ettiği gözlenmiştir.

Classification of Electrical Appliance Devices Using –Bi-LSTM and Genetic Algorithm

Abstract

Keywords

LSTM;Bi-LSTM;
Electrical Appliances;
Current/voltage;
Genetic Algorithm.

Since the current/voltage characteristics of each electrical appliance are different, the power consumption of devices are specific. Therefore, determining the type of electrical appliance connected to the network is crucial for the detection and correction of the device based harmonics. In this study, the classification of electrical appliances is carried out on the "ACS-F2 Electrical Appliances Dataset" using different deep learning algorithms. Although there are 225 devices for 15 different classes in the ACS-F2 data set, the number of classes in the data set has been reduced to 14 with the pre-processing step. The LSTM, FeedForwardNet, Bi-LSTM and Ga+Bi-LSTM, models are then built to classify electrical appliances. It is observed that the GA+Bi-LSTM classifier, which has %94 classification accuracy, overcomes among the models.

© Afyon Kocatepe Üniversitesi

1. Giriş

Günümüzde elektrik enerjisine bağlı olan cihaz sayısı ve buna bağlı olarak da enerji ihtiyacı gittikçe artmaktadır. Bu cihazlar arasında elektrikli ev aletleri daha az enerji tüketmesine karşın toplamda daha fazla sayıda olduğundan dolayı, bu cihazların enerji karakteristiğinin gözlenmesi ve bu karakteristiğin belirlenmesi enerji tüketimlerinin azaltılmasında önemlidir. Bu sayede hem enerji

verimliliği (cihazların normalden fazla kullanılmasının tespiti) sağlanabileceği gibi hem de cihazlardaki arızaların tespiti sağlanabilir (Howell 2007, Patel et al. 2007, Hart 1997, Kuruppu and. Kulatunga 2012).

Elektrikli cihazların yük karakteristikleri V_{rm}, I_{rms}, aktif güç, reaktif güç ve görünür güç gibi parametreler kullanılarak belirlenmektedir. Bu bağlamda faz açısı elektrikli cihazların yük

karakteristiğinin belirlenmesinde en önemli parametre olarak karşımıza çıkmaktadır. Elektrikli cihazlar, faz açısına göre rezistif, kapasitif ve endüktif olarak sınıflandırılmaktadır. Su ısıtıcıları gibi elektrikli ısıtıcılar rezistif bir davranış gösterirken, indüksiyon motorları ve floresan lambalar endüktif, tekrar şarj edilebilen bataryalar da kapasitif davranış göstermektedir. Birçok durumda elektrikli cihazlar aktif ve reaktif güç değerleri kullanılarak birbirinden ayrılabilir (Armel et al. 2013). Ölçümler genellikle binaların/ evlerin giriş noktasından (ana giriş noktasından) bir akıllı sayaç yardımıyla yapılmasına karşın bu yöntem her bir cihazın akım gerilim karakteristiği farklı olduğundan dolayı doğru bir yöntem değildir. Bunun yerine her bir cihazın akım/gerilim karakteristiğinin belirlenmesi gerekmektedir.

Ridi ve arkadaşları yaptıkları çalışmada elektrikli ev aletlerinin karakteristiklerini belirlemek için birçok farklı elektrikli ev aletlerinden veri toplayıp elde ettikleri veri setini ACS-F2 ismiyle paylaşmışlardır (Ridi et al. 2014). Yazarlar veri setini elde ederken cihazların normal çalışma koşullarında olmasına özellikle dikkat etmişlerdir. Yazarlar verileri cihazlardan toplarken aşağıdaki rutinleri özellikle uygulamışlardır.

Mekanik Fanlar : Fanlar üç farklı çalışma moduna iken (kesim, bekleme modu ve çalışma modu) veriler toplanmıştır. Çalışma modunda veriler toplanırken mekanik fanların bütün hızlarda dönmesi sağlanmıştır.

Elektrikli Su Isıtıcıları : İçerisinde farklı seviyede su eklenerek çalıştırılmış ve veriler bu şekilde toplanmıştır.

Monitörler: Monitörlerden veriler toplanırken her biri farklı cihazlara (masa üstü bilgisayar, laptop) bağlanmıştır. Farklı cihazlardaki bağlı olan monitörlerden veriler toplanırken, kullanıcıların doğal davranışları istenmiş ve de veriler doğal bir şekilde toplanmıştır. Buna ek olarak ekran parlaklığının değişiminin etkisi de alınan verisetinde incelenmiştir.

Traş Makineleri: Bu cihazlardan veriler toplanırken bütün modlar (çalışma modu, bekleme modu) göz önüne alınmıştır. Çalışma modunda ise bıçakların bütün dönüş hızları ve tork değerleri göz önüne alınarak veriler toplanmıştır.

Emir ve arkadaşları yaptıkları çalışmada ACS-F2 veriseti üzerinde rastgele orman, Bayes sınıflandırıcı ve de Hoeffding ağacı sınıflandırma algoritması olmak üzere üç farklı sınıflandırma algoritması denemişlerdir (Emir et. al 2016). Yazarlar öz-nitelik olarak gücün, faz açısının, akımın istatistiksel parametrelerini (standart sapma, maximum ve minimum değerleri, basıklık, çarpıklık) kullanmışlardır. Yazarlar yapılan çalışmada eğitim veri seti üzerinde %93.67 doğruluk elde etmelerine karşın, test veri seti üzerinde elde ettikleri doğruluk %68.67 seviyesinde kalmıştır. Samira ve arkadaşları yaptıkları çalışmada ACS-F2 veri setinin sınıflandırılması için Rastgele orman, k-NN, SVM ve GMM modellerini kullanmışlardır (Samira and Rammohan 2018). Yazarlar dört farklı senaryoya göre her bir modelin elde ettiği doğruluk seviyesini karşılaştırmışlardır. Yazarlar rastgele orman algoritmasının %75.65 ile en yüksek doğruluğu elde ettiğini belirtmişlerdir.

Long Short Term Memory (LSTM) yapıları Recurrent Neural Network(RNN)'lerin özelleştirilmiş bir hali olarak ilk olarak 1997 yılında Hochreiter tarafından önerilmiştir (Hochreiter and Jurrgen 1997). O günden sonra LSTM'ler esnek yapısı sayesinde zaman serisi tahminleme başta olmak üzere, birçok farklı sınıflandırma algoritmasında başarılı bir şekilde kullanılmıştır (Pulwer and Lyu, 2017). Fakat LSTM'lerin parametrelerinin ayarlanması da önemli bir sorun olarak karşımıza çıkmaktadır. Genellikle probleme göre değişmekle beraber, bu parametrelerin optimal bir şekilde belirlenmesi bir optimizasyon problemi olarak karşımıza çıkmaktadır. Çift yönlü LSTM'ler (Bi-LSTM) de başarı üstünlükleri nedeniyle birçok farklı alanda kullanılmıştır (Jiajun Li et. al 2022).

Bu çalışma kapsamında, ACS-F2 veri seti kullanılarak elektrikli ev aletinin sınıflandırılması gerçekleştirilmiştir. Bu amaçla, öncelikli olarak veriler üzerinde bir ön-eleme işlemi gerçekleştirilmiştir. Bu ön-eleme işlemi sonucunda hatalı ölçüm olduğu belirlenen veriler elenerek veri setinden silinmiştir. Sonrasında, genetik algoritma tarafından optimize edilmiş Bi-LSTM, Bi-LSTM, LSTM ve Feed-Forward NET olmak üzere dört farklı sınıflandırıcı kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Geliştirilen Genetik algoritma

destekli Bi-LSTM modelinin her bir sınıf için %85 üzerinde bir doğruluk elde ettiği gözlenmiştir. Öte yandan diğer sınıflandırıcıların ortalama doğruluğu sırasıyla %81, %77, %33 olarak hesaplanmıştır. Geliştirilen GA destekli Bi-LSTM sınıflandırıcının elde ettiği sonuçlar, istatistiksel olarak literatürde yapılan diğer çalışmalarla karşılaştırılmış ve literatürde yapılan çalışmalardan daha üstün olduğu gözlenmiştir. Geliştirilen bütün algoritmalar GPU desteği olan bir bilgisayar ile MATLAB ortamında geliştirilmiştir.

2. Materyal ve Metot

2.1 Veri Kümesi ve Ön İşlemler

Bu çalışma kapsamında “ACS-F2 Elektrikli Ev Aletleri Veri Seti”den elde edilmiştir. ACS-F2 veri seti verilen bağlantıdan .xml uzantılı olarak indirilebilir. (<https://icosys.ch/acs-f2>) Veri setinde toplamda 15 farklı sınıf için 225 farklı cihazdan örnekler toplanmıştır. Bu 15 sınıf sırasıyla: Kahve makinası, bilgisayar, vantilatör, buzdolabı, ses sistemi, su ısıtıcı, floresan lamba, akkor lamba, laptop, mikrodalga, cep telefonu, monitör, yazıcı, traş makinası, televizyon olarak belirlenmiştir.

Bu veri setinde geliştiriciler tarafından temel frekans (fe), faz açısı, görünür güç (S), reaktif güç (Q), akım ve gerilimlerin RMS değerleri cihazlardan 100 ms’lik periyotlarla toplanmıştır. Sonrasında bu toplanan değerler ile altı elemanlı öznitelik vektörleri oluşturulmuştur. Bu çalışma kapsamında .xml formatındaki bu veriler bir matris haline getirilerek MATLAB ortamında işlenmiştir. Matrisin son sütunlarında cihazın gerçek sınıf değeri tutulmuştur. Bu sayede elde edilen veri setinin boyutları 161108*7 olmuştur.

Veriseti üzerinde gerçekleştirilen incelemede ACS-F2 veri setinin birinci sütunda bulunan temel frekans değerinin alınan örnekler boyunca sabit kaldığı gözlenmiştir. Bu nedenle, bu frekans değeri hem sınıflandırıcıların performansını azaltmakta hem de hesapsal yükü arttırdığından dolayı öznitelik vektöründen çıkartılmıştır. Yine yapılan ön incelemelerde bazı hatalı ölçüm verileri tespit edilmiştir. Örneğin akımın (I) RMS değeri “0” olarak ölçülmüş olmasına karşın, bir görünür güç değerinin ($S=V_{rms} * I_{rms}$) hesaplandığı gözlenmiştir. Bu nedenle hatalı ölçümleri tespit eden ve bu değerleri

öznitelik veri kümesinden elimine eden bir rutin geliştirilmiştir.

Yapılan ön işlemler sonucunda 6 numaralı sınıfa yani su ısıtıcısına ait veri kalmadığı gözlenmiştir. Bu nedenle sınıf sayısı 15’den 14’e düşürülmüştür. Bu ön işlemler sonucunda veri matrisinin boyutu 161108x6’dan 90495x5’e düşürülmüştür.

Detaylı bir şekilde açıklandığı gibi veri seti üzerinde gerçekleştirilen ön işlemlerden sonra veri setindeki bazı elemanlar elendiğinden dolayı her bir sınıftan eşit sayıda veri bulunmamaktadır. Bu nedenle, ön işlemlerden sonra, eğitim ve test verilerinin oranlarını belirlemek için her bir sınıfta toplamda kaç adet verinin kaldığı tespit edilmiş ve veri setinde dengesizlik olmaması için veri sayısının en az olduğu sınıfa göre eğitim ve test verilerinin sayısı belirlenmiştir. Tablo1’de yapılan ön işlemlerden sonra her bir sınıf için kalan veri sayısı verilmiştir.

Tablo 1. Ön eleme sonrasında sınıflarda kalan eleman sayıları

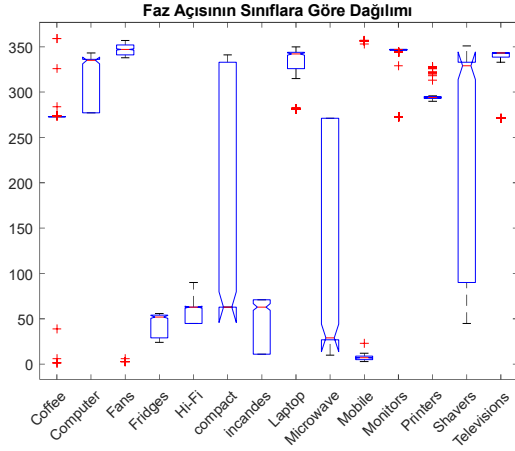
Sınıf	Eleman Sayısı
1	7580
2	10037
3	4466
4	4882
5	9789
6	0
7	6159
8	1776
9	10151
10	6020
11	625
12	9605
13	7074
14	1950
15	10391

Kalan veriler incelendiğinde model geliştirilirken, veri setinde dengesizlik olmaması için tüm sınıflardan eşit sayıda örnekler alınarak eğitim ve test kümeleri oluşturulmuştur. Tablo 1 incelendiğinde en az verinin 11 numaralı sınıfa yani cep telefonuna ait olduğu gözlenmiştir. Bu nedenle kullanılacak olan test ve eğitim verilerinin sayısı bu değerden az olmalıdır. Bu nedenle eğitim verisi

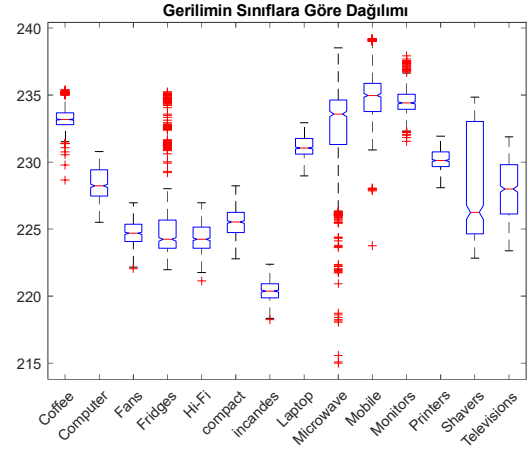
olarak 625 elemanın %80'i yani, her bir sınıftan rastgele olarak 500 eleman seçilerek eğitim gerçekleştirilmiştir. Hesapsal yükü azaltmak amacıyla, kalan 125 elemandan rastgele seçilen 50

tanesi geliştirilen modellerin test edilmesinde kullanılmıştır.

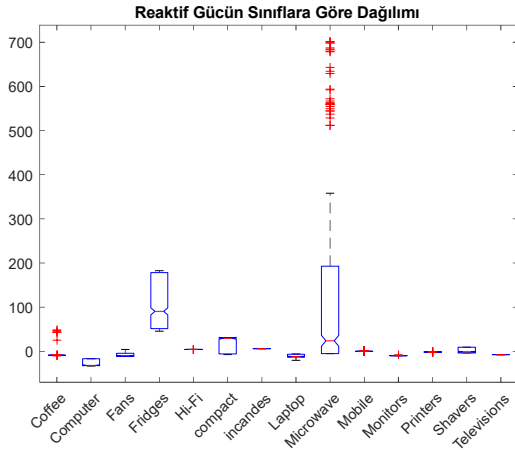
Her bir sınıfa göre öznitelikleri dağılımları kullanılarak Şekil 1'de verilmiştir.



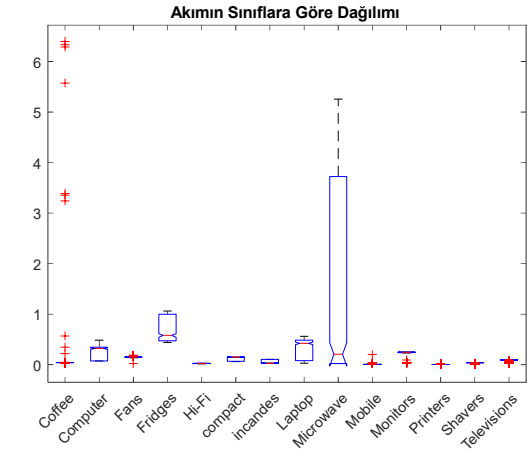
(a)



(b)



(c)



(d)

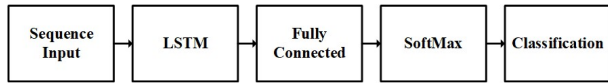
Şekil 1: Özniteliklerin sınıflara göre dağılımları a: Faz açısı, b: Vrms, c: Q, d:Irms

2.2 LSTM tabanlı sınıflandırma

Long short term memory (LSTM) ilk olarak 1997 yılında Hochreiter & Schmidhuber tarafından geliştirilmiş bir Recurrent Neural Network (RNN) modelidir. RNN modellerinde de tutulan geçmiş bilgilere ait gradyenler eğitim sırasında yok olabilmektedir. Bu problem de ağlar arasındaki katsayıların çok büyük veya çok küçük olmasına neden olur. Bu sorunların üstesinden gelmek için tekrarlayan bir ağ yapısı olan LSTM kullanılmaktadır. LSTM, gizli katmanlarda tutulan bilgilerin hangi çıktıya ve gelecekte hangi gizli katmana gideceğini kontrol etmek için ek kapılar

kullanır. LSTM, hücre durumunu kullanılan bu ek kapılarla bir sonraki hücreye bilgi ekleme veya silme şeklinde iletebilirler.

Şekil 2'de bu çalışma kapsamında kullanılan LSTM yapısı gösterilmiştir. LSTM'ler esnek yapısı sayesinde görüntü sınıflandırma başta olmak üzere zaman serisi tahminleme, sınıflandırma gibi birçok farklı alanda başarılı bir şekilde uygulanmıştır. Bu çalışma kapsamında da elimizde sinyallerden çıkartılmış olduğundan, LSTM'ler sınıflandırma amaçlı olarak kullanılmıştır.



Şekil2: Kullanılan LSTM yapısı

Bu çalışma kapsamında LSTM'ler sinyal sınıflandırıcı olarak kullanılacağından, giriş olarak öznitelik vektörü verilmiştir. Bu nedenle giriş katmanında sequence input bulunmaktadır. Benzer şekilde Bi-Directional LSTM'de de benzer şekilde bir yapı kurulmuştur. Geliştirilen LSTM tabanlı sınıflandırıcıların parametreleri Tablo 2'de verilmiştir.

Tablo 2: LSTM sınıflandırıcı parametreleri

Parametre	Değeri
MiniBatch	5
Epoch	5
Gizli Katman Sayısı	100
Aktivasyon Fonksiyonu	SoftMax

2.3 Genetik Algoritma

Sezgisel algoritmalarından biri olan, genetik algoritma temelde Darwin'in evrim teorisine dayanmaktadır. Bu nedenle, gen, doğal seçim, mutasyon gibi kavramlar genetik algoritmanın temellerini oluşturmaktadır. Özellikle çaprazlama ve mutasyon işlemleri ile lokal ekstremum noktalarından kurtulabilmesi genetik algoritmanın en büyük avantajlarından biridir. Bu sayede, genetik algoritma birçok farklı optimizasyon çalışmalarında başarılı bir şekilde kullanılmıştır (Dehzen and Kail 2008). Özellikle makine öğrenmesi ve de derin öğrenme modellerinde modelin parametrelerinin optimizasyonunda sıklıkla kullanılmaktadır.

Bu çalışma kapsamında da sınıflandırmada kullanılan Bi-LSTM derin öğrenme modelinin hiperparametrelerinin optimizasyonu sağlanmıştır. Bu çalışma kapsamında batch size, epoch size, gizli katman sayısı ve aktivasyon fonksiyonu optimize edilmiştir. Sınıflandırıcının doğruluk değeri genetik algoritmanın uygunluk fonksiyonu olarak belirlenmiştir. Genetik algoritma sezgisel bir yöntem olduğundan her bir çalışma sırasında farklı bir çözüm elde edebileceğinden dolayı, geliştirilen algoritma 100 defa çalıştırılarak elde edilen

sonuçların ortalaması alınmıştır. Yapılan çalışma sonucunda optimal parametrele Epoch, MiniBatch, Gizli katman boyutu, ve aktivasyon fonksiyonu için sırasıyla, 5,3,100 ve SoftMax olarak bulunmuştur.

3. Bulgular

Bu çalışma kapsamında kullanılan veriseti üzerinde geliştirilen her bir sınıflandırıcının sınıflandırma doğruluğu Tablo 3'de verilmiştir. Tablo 3 incelendiğinde Geliştirilen GA tarafından optimize edilmiş Bi-LSTM Modelinin monitör hariç kalan sınıflar için en yüksek sınıflandırma doğruluğuna sahip olduğu gözlenmiştir. Her bir sınıf için sınıflandırma doğruluğunun %76'in üzerinde ve ortalama doğruluğun da %94 olduğu gözlenmiş olup, bu değerlerin de literatürde yapılan çalışmalardan daha üstün olduğu Tablo 4'de diğer yapılan çalışmalarla karşılaştırılarak gösterilmiştir.

Tablo 3: Sınıflandırıcıların doğrulukları

Sınıf	GA+BiLSTM	BiLSTM	LSTM	FeedForwardNet
Kahve makinası	1	0.98	0.98	0.22
Bilgisayar	1	1	1	0.12
Vantilatör	1	1	0.28	0.15
Buzdolabı	1	1	1	0.27
Sessistemi	0.98	0.98	0.98	0.37
Floresan lamba	0.98	0.98	0.68	0.23
Akkor lamba	1	0.98	1	0.28
Laptop	0.76	0.56	0.52	0.05
Mikrodalg a	0.86	0.74	0.64	0.09
Cep telefonu	1	1	0.98	0.33
Monitör	0.78	0.88	0.88	0.34
Yazıcı	1.	1.	0.98	0.41
Tıraş makinası	0.88	0.8	0.96	0.35
Televizyon	0.95	0.9	1	0.29
Ortalama	0.94	0.91	0.84	0.25

Yapılan çalışma kapsamında geliştirilen GA destekli Bi-LSTM derin öğrenme modelinin elde ettiği başarı oranı Tablo 4'de literatürde bulunan diğer

yöntemlerle karşılaştırılmıştır. **Tablo 4:** Geliştirilen Modelin Literatürdeki yöntemlerle karşılaştırılması

Model	Ortalama doğruluk
Emir	%67
Ridi	%73
Samira	%68
Önerile Model (GA+BiLSTM)	%94

4. Sonuç

Bu çalışmada, kapsamında elektrikli ev aletlerinin şebekeden çektikleri akım/gerilim karakteristiğine göre sınıflandırılması için metotlar geliştirilmiştir. Geliştirilen metotlar ACS-F2 veriseti üzerinde denenmiştir. Geliştirilen yöntemlerden genetik algoritma destekli Bi-LSTM modelinin diğer yöntemlerden çok daha yüksek sınıflandırma doğruluğu(%94) elde ettiği gözlenmiştir. Çalışma kapsamında önerilen modelin literatürde yapılan çalışmalardan çok daha başarılı sonuçlar elde ettiği gözlenmiştir.

5. Kaynaklar

- Howell E.K. 2007. How switches produce electrical noise," IEEE Transactions on Electromagnetic Compatibility, no. **3**, pp. 162–170,
- Patel S.N., Robertson T., Kientz J.A., Reynolds M.S. Abowd., G.D, 2007. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line, 9th international conference on Ubiquitous computing, Springer-Verlag, pp. 271–288.
- Hart G. W., 1992, Nonintrusive appliance load monitoring," Proceedings of the IEEE, vol. **80**, pp. 1870-1891,

Kuruppu S., Kulatunga N. A., 2012.Smart meter based non-intrusive appliance detection algorithm for local real time feedback purposes," in Innovative Smart Grid Technologies-Asia (ISGT Asia), 2012 IEEE, , pp. 1-5.

Armel K. C., Gupta A., Shrimali G., Albert A., 2013.Is disaggregation the holy grail of energy efficiency? The case of electricity," Energy Policy, **52**, pp. 213-234,

Ridi A., Gisler C., Hennebert J., 2014. A survey on intrusive load monitoring for appliance recognition, 22nd International Conference on Pattern Recognition (ICPR),

Salihagic E., Kevric J., Dogru N., "Classification of ON-OFF states of Appliance Consumption Signatures" XI International Symposium on Telecommunications. 2016.

Samira G., Rammohan M., 2018. Significance of Classifier and Feature Selection in Automatic Identification of Electrical Appliances IEEE International Conference on Systems, Man, and Cybernetics,

Hochreiter S., Jurgen S., 1997. Long Short Term Memory N. Journal of Neural Computation, **9(1)**, 1735-1780.

Pulwer A., Liu S., "LSTM with working memory" N. International Joint Conference on Neural Networks (IJCNN)

Jiajun L., Huazhu S., Jun L., Kaituo M., "Comprehensive analysis and classification of natural language questions based on Bi-LSTM-CRF" IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)

Dezhen Z., Kai Y., 2008. Genetic Algorithm Based Optimization for AdaBoost, International Conference on Computer Science and Software Engineering,